Politecnico di Torino

PhD thesis in Cultural Heritage
XXVIII cycle
ICAR06

Spatial ontologies for architectural heritage

Francesca Noardo
2016

Coordinator: prof. C. Roggero
Politecnico di Torino – DIST
Interuniversity Department of regional and Urban Studies and planning

Tutors: prof. A. Spanò
Politecnico di Torino – DAD
Department of Architecture and Design

prof. A. Lingua
Politecnico di Torino – DIATI
Department of Environment, Land and Infrastructure Engineering

There is an obvious element of beauty in engineering artefacts in which aesthetics is a criterion for quality. For example, in architecture, it is considered since Roman times that a good building should satisfy the three principles of firmitas, utilitas, venustas, i.e., durability, utility and beauty.

"Is there beauty in ontologies?"
(d'Aquin, Gangemi, 2011)

Ai miei bimbi sperduti

ABSTRACT

Informatics and artificial intelligence have generated new requirements for digital archiving, information, and documentation. Semantic interoperability has become fundamental for the management and sharing of information. The constraints to data interpretation enable both database interoperability, for data and schemas sharing and reuse, and information retrieval in large datasets. Another challenging issue is the exploitation of automated reasoning possibilities. The solution is the use of domain ontologies as a reference for data modelling in information systems. The architectural heritage (AH) domain is considered in this thesis. The documentation in this field, particularly complex and multifaceted, is well-known to be critical for the preservation, knowledge, and promotion of the monuments. For these reasons, digital inventories, also exploiting standards and new semantic technologies, are developed by international organisations (Getty Institute, ONU, European Union). Geometric and geographic information is essential part of a monument. It is composed by a number of aspects (spatial, topological, and mereological relations; accuracy; multi-scale representation; time; etc.). Currently, geomatics permits the obtaining of very accurate and dense 3D models (possibly enriched with textures) and derived products, in both raster and vector format. Many standards were published for the geographic field or in the cultural heritage domain. However, the first ones are limited in the foreseen representation scales (the maximum is achieved by OGC CityGML), and the semantic values do not consider the full semantic richness of AH. The second ones (especially the core ontology CIDOC – CRM, the Conceptual Reference Model of the Documentation Commettee of the International Council of Museums) were employed to document museums’ objects. Even if it was recently extended to standing buildings and a spatial extension was included, the integration of complex 3D models has not yet been achieved. In this thesis, the aspects (especially spatial issues) to consider in the documentation of monuments are analysed. In the light of them, the OGC CityGML is extended for the management of AH complexity. An approach ‘from the landscape to the detail’ is used, for considering the monument in a wider system, which is essential for analysis and reasoning about such complex objects. An implementation test is conducted on a case study, preferring open source applications.

TABLE OF CONTENT

ABSTRACT III

INTRODUCTION 1

The architectural heritage documentation as a ressource 1

ontologies: interoperability and automated reasoning 9

SECTION 1 – THE RESEARCH OBJECTIVES 13

1 The characteristics of architectural heritage documentation 15

1.1 The thematic side 16
1.2 The geometric side: from cartography to architectural large scale 17
1.3 Further elements for spatial knowledge 19
1.4 Time 27
1.5 An example: an architectural part and reasoning mechanisms 27

2 The life of architectural heritage information 29

2.1 The production 29
2.2 The management 44
2.3 The analysis 60
2.4 The communication 65

3 GIS for large-scale architecture information: state of the art 68

3.1 GIS for general documentation aims 69
3.2 GIS as a support for a preservation project 75
3.3 Stratigraphic Analysis in GIS 80
3.4 Final comments 85

4 CH Digital inventories and 3D-models web archives 86

4.1 The Getty institute 86
4.2 The Smithsonian Institution 87
4.3 The Smithsonian X 3D 88
4.4 CyArk 89

SECTION 2 – THE METHOD 91

5 The domain ontologies 93

5.1 The philosophical ontology 93
5.2 Applied ontologies 94

6 Interoperability exigencies: standardisation 108

6.1 International Organization for Standardization (ISO) 109
6.2 Cultural Heritage Standards 110
6.3 Standardisation in the field of digital geographic information 119

7 Smart communities and the web evolution for information retrieval 135

7.1 The Semantic Web 136

7.2 The Linked Open Data ... 139
7.3 The concept of Uniform Resource Identifier (URI) ... 140
7.4 W3C (World Wide Web Consortium) ... 141
7.5 Some Semantic Web tools for data meaning management ... 144
7.6 OPEN STAND - The Modern Paradigm for Standards ... 146
7.7 An inherent issue: the open source community and the open source software ... 147
7.8 Domain ontologies in the Semantic Web ... 148
7.9 OGC answers to the Semantic Web proposals ... 167

SECTION 3 – THE APPLICATION OF INTEROPERABILITY TECHNOLOGIES TO ARCHITECTURAL HERITAGE ... 171

8 A data model for architectural heritage management and representation ... 173

8.1 The CityGML Cultural Heritage Application Domain Extension ... 173

9 Managing a 3D model in Citygml CHADE: Preparing and filling in the 3D data ... 182

9.1 The case study ... 182
9.2 Reduction of the computational intensiveness of the products ... 184
9.3 Semantic Segmentation ... 184
9.4 Technical issues for editing of models ... 189
9.5 The translation to CityGML CHADE format ... 189

10 Results: surfing the archive. Visualisation and analysis in a graphical interface ... 193

11 An alternative use of the CityGML CHADE model: the generation of an SQL database 200

12 The further step: the translation of the GML data to Linked Open Data ... 201

SECTION 4 – CONCLUSIONS AND PERSPECTIVES ... 203

13 A system towards documentation intelligence ... 205

REFERENCES ... 209

ANNEXES ... 227

INTRODUCTION

THE ARCHITECTURAL HERITAGE DOCUMENTATION AS A RESSOURCE

'[…] it has become more and more necessary not only to promote common values in the management and enhancement of the cultural heritage but also through cultural co-operation to provide the common mechanisms that allow us to preserve, celebrate, and enrich that heritage for the benefit of those who follow. Accurate documentation is fundamental to this task.'

José María Ballester
Head of the Cultural Heritage Department
Council of Europe[1]

Since the early documents arose from the awareness of the importance of cultural heritage for humanity, this documentation is considered as a fundamental requirement for the cultural heritage to be known, preserved, and promoted.

The first important international document for the preservation of monuments, the *Athens Charter*, in 1931 states:

'[…] art. VIII: La Conferenza emette il voto: 1. che i vari Stati, ovvero le istituzioni in essi create o riconosciute competenti a questo fine, pubblichino un inventario dei monumenti storici nazionali accompagnato da fotografie e da notizie; 2. che ogni Stato crei un archivio ove siano conservati i documenti relativi ai propri monumenti storici;[…].

art. X: La Conferenza, profondamente convinta che la migliore garanzia di conservazione dei monumenti e delle opere d'arte venga dall'affetto e dal rispetto del popolo […]'[2]

Some important concepts are already present in these words: the necessity of filling in cultural heritage public inventories and archives, with multi-media materials (including photographs and, generally, information). Moreover, the importance of the people's consideration towards monuments as a guarantee for their preservation has already been affirmed. This can be primarily encouraged through knowledge and communication policies, as asserted also by the recent rules of cultural heritage promotion (Pickard, 2002).

It is clear how digital archiving technologies can improve the performances of inventorying multi-media documents and how new technologies of communication, including the internet, can presently empower the public knowledge and promotion of cultural heritage.

Similar ideas were stated in the following year, 1932, in the *Carta Italiana del Restauro* (Italian Preservation Charter), in which the importance of the monument for the citizens and, enhancing the application field to wider scales, the urban spirit, are stated. In the document, besides the artistic and

[1] http://archives.icom.museum/objectid/heritage/fore.html

[2] The translation of these words is: '[…] art. VIII: The Conference makes the vow: 1. that the various States or the institutions they created or recognized competent for this purpose, publish an inventory of national monuments, along with photos and news; 2. that each State should create an archive where the documents related to its historic monuments are stored;[...].
art. X: The Conference, firmly convinced that the best guarantee of preservation of monuments and works of art is by the affection and respect of the people [...]'.

historic values, the practical function of the monument is considered, as a further element of the complexity of such objects. Finally, also in the Italian Charter, the redaction and publication of a document about the knowledge about monuments is deemed as a necessity, paying attention also to the recording of the possible stratifications that can emerge during works (such as restoration operations).

'Art. XI: che come nello scavo, così nel restauro dei monumenti sia condizione essenziale e tassativa, che una documentazione precisa accompagni i lavori mediante relazioni analitiche raccolte in un giornale del restauro e illustrate da disegni e da fotografie, sicché tutti gli elementi determinanti nella struttura e nella forma del monumento, tutte le fasi delle opere di ricomposizione, di liberazione, di completamento, risultino acquisite in modo permanente e sicuro. […]'

'[…] Conclusioni, voto 3: che sia fatto obbligo della compilazione e della conservazione metodica dei suddetti giornali del restauro, e che possibilmente dei dati e delle notizie analitiche da quelli risultanti si curi la pubblicazione scientifica […]'[3] *(Carta Italiana del Restauro, 1932).*

The following Preservation Charters reaffirm and articulate the same concepts defining increasingly many multifaceted characteristics regarding architectural heritage. They consist in: its value as a testimony of a civilization, a meaningful evolution or historic event; the inclusion of elements of sculpture, painting, or decoration as an integral part of the monument; the importance of the graphic and photographic survey including metric information; the importance of the static conditions (*Carta di Venezia per il restauro e la conservazione di monumenti e siti*, Venice Charter for the preservation of monuments and sites, 1964); the possibility to consider the complexes of buildings as an architectural heritage item (*Carta Italiana del Restauro*, Italian Restauration Charter, 1972); and the same considerations are applied to urban aggregations (*Carta della conservazione e del restauro degli oggetti d'arte e di cultura*, Charter for the preservation and restauration of art and culture objects, 1987). A synthesis of the recording principles stated by the international charters is present in (Letellier et al., 2007 - Appendix H).

The architectural heritage documentation is also effectively called 'recording'. The definition given by the ICOMOS (International Council on Monuments and Sites) assembly follows.

*'**Recording** is the capture of information which describes the physical configuration, condition and use of monuments, groups of buildings and sites, at points in time, and it is an essential part of the conservation process. **Records** of monuments, groups of buildings and sites may include tangible as well as intangible evidence, and constitute a part of the documentation that can contribute to an understanding of the heritage and its related values.'*[4]

In the same text, the principle is affirmed that '*The format of the records should be standardized, and records should be indexed wherever possible to facilitate the exchange and retrieval of information at a local, national or international level*'. Moreover, the necessity to consider an

[3] 'Art. XI: as in the (arqueological) excavation, so in the restoration of monuments, be it essential and mandatory, that precise documentation accompany the work by means of analytical reports collected in a restoration paper and illustrated with drawings and photographs, so that all crucial factors in the structure and in the form of the monument, all stages of reconstruction works, liberation, completion, are acquired permanently.[...]'

'[...] Conclusions, vow 3: that be it required the filling-in and methodic conservation of those restoration papers, and possibly of data and analytical information, from those resulting, scientific publication be cared [...]'

[4] Principles for the recording of monuments, groups of buildings and sites, 11th ICOMOS General Assembly, Sofia, Bulgaria, from 5 to 9 October 1996.

appropriate up-to-date technology and to make the location of the data known is stated in the same section, 'Management, dissemination and sharing of the records'.

In more recent years, an explicit and inclusive principle regarding the information (in particular, the principle 2: information sources) has been stated in the 'Ename Charter' or, better, 'The ICOMOS Charter for the Interpretation and Preservation of Cultural Heritage Sites in 2007'[5]. Here, important terms are defined, such as, in addition to the concept of 'Cultural Heritage Site', the object of the Charter regulation ('interpretation', or 'presentation') which are the objectives at which the research presented in this thesis aims.

*"**Cultural Heritage Site** refers to a place, locality, natural landscape, settlement area, architectural complex, archaeological site, or standing structure that is recognized and often legally protected as a place of historical and cultural significance."*

***Interpretation** refers to the full range of potential activities intended to heighten public awareness and enhance understanding of cultural heritage site […].*

***Presentation** more specifically denotes the carefully planned communication of interpretive content through the arrangement of interpretive information, physical access, and interpretive infrastructure at a cultural heritage site […].*

***Interpretive infrastructure** refers to physical installations, facilities, and areas at, or connected with a cultural heritage site that may be specifically utilised for the purposes of interpretation and presentation including those supporting interpretation via new and existing technologies."*[6]

The objectives of the Ename Charter (cited in the italic text below) regard the understanding and the interpretation of the items defined as 'cultural sites' (including architectural heritage) through instruments that permit the documentation of the sites and include all that elements that constitute the essence of a cultural heritage item. The communication of the built knowledge fosters the appropriate appreciation of the sites and, consequently, confers them the proper respect and the awareness of the importance of their preservation.

*"1. **Facilitate understanding and appreciation** of cultural heritage sites and foster public awareness and engagement in the need for their protection and conservation.*

*2. **Communicate the meaning** of cultural heritage sites to a range of audiences through careful, documented recognition of significance, through accepted scientific and scholarly methods as well as from living cultural traditions.*

*3. **Safeguard the tangible and intangible values** of cultural heritage sites in their natural and cultural settings and social contexts.*

[5] http://www.enamecharter.org/downloads/ ICOMOS_Interpretation_Charter_EN_10-04-07.pdf

[6] The italic text cites the "Ename Charter"

*4. **Respect the authenticity** of cultural heritage sites, by communicating the significance of their historic fabric and cultural values and protecting them from the adverse impact of intrusive interpretive infrastructure, visitor pressure, inaccurate or inappropriate interpretation.*

*5. **Contribute to the sustainable conservation** of cultural heritage sites, through promoting public understanding of, and participation in, ongoing conservation efforts, ensuring long-term maintenance of the interpretive infrastructure and regular review of its interpretive contents.*

*6. **Encourage inclusiveness** in the interpretation of cultural heritage sites, by facilitating the involvement of stakeholders and associated communities in the development and implementation of interpretive programmes.*

*7. **Develop technical and professional guidelines** for heritage interpretation and presentation, including technologies, research, and training. Such guidelines must be appropriate and sustainable in their social contexts."*[7]

As summarized by (Stylianidis et al., 2011) the information plays an essential role in safeguarding cultural heritage, since it permits:

- *Understanding of significance and integrity of heritage places;*
- *Transmission of this understanding thus increasing public awareness;*
- *Informing decisions for conservation;*
- *Creating records for posterity in case of destruction;*
- *Providing timely and sufficient information for preventive maintenance.*[8]

RecorDIM: a reference for cultural heritage documentation

[International community – 2003-2007]

The RecorDIM initiative reasons and produced guidelines can effectively summarize the necessities and principles of cultural heritage documentation defined by all the previous international documents.

The RecorDIM (Recording, Documentation and Information Management) initiative was promoted by most of the organizations dealing with cultural heritage at the international level: ICOMOS (International Council on Monuments and Sites), CIPA (Comité International de Photogrammétrie Architecturale, that is 'International Committee about the Architectural Photogrammetry', but is today imparted with a wider meaning 'International Committee for documentation of Cultural Heritage'), Heritage Documentation, English Heritage, World Monuments Fund, Heritage Conservation Directorate of Public Works & Government Services Canada, ICCROM (International Centre for the Study of the Preservation and Restoration of Cultural Property), India National Trust for Art and Cultural Heritage, Université du Québec à Montréal, Comité International des Itinéraires

[7] The italic text cites the "Ename Charter"
[8] The italic text cites (Stylianidis et al., 2011)

Culturels (ICOMOS-ISC), Comité International de Formation (ICOMOS-ISC) Politecnico di Torino, the University of Pennsylvania, and the World Heritage Center.

The aim of this project was to bridge the gaps between information users (including researchers, conservation specialists of all trades, project managers, and planners) and information providers (including photographers, heritage recorders, photogrammetrists, and surveyors) following the contents of the ICOMOS document 'Principles for the recording of monuments, groups of buildings and sites' (1996) (Eppich, Chabbi, 2007).

In the project, a series of guiding principles are stated, encouraging the sharing of the information through every means available (internet included), but a specific standard to be used is not defined for the recording, management, and communication of the data.

Guiding principles RecorDIM:

How should records be kept and identified? Original records of heritage places must use standardized formats, be preserved in a safe and accessible place, be backed up, and, in the case of digital records, be regularly migrated to the most current versions of software and support. Although costly, archiving hard copies of digital records is also a recommended practice. For records to be easily retrieved and managed, a unique identifier is required, such as standard longitude and latitude coordinates that define the location of a heritage place.

The preamble recognizes three related points:

- *The inherent value of the cultural heritage, that is, the "unique expression of human achievement"*
- *The risk to which that heritage is continually exposed*
- *The recognition of recording as one of the principal means to improve understanding of the values associated with cultural heritage […]*

Recording is seen as essential for the following reasons.

- *Recording enhances understanding of cultural heritage.*
- *Recording promotes the involvement of the public.*
- *Recording improves the quality of management decision-making affecting cultural heritage at all levels, including decisions concerning appropriate use.*
- *Recording helps ensure that planned interventions respect the defined qualities and characteristics of heritage places.*
- *Recording provides a permanent record of cultural heritage prior to change, planned or unplanned.*[9]

[9] Italic text cites (Letellier et al., 2007).

An enhanced and upgraded role for cultural heritage

In a very recent European and international framework, new problems have arisen or have acquired a new relevance and urgency. Among these are included economic crisis, immigration problems, new terrorism actions, unfavourable effects of globalization, increasing social and political disengagement of young people, instances of unsustainable economic development, climate changes, and consequent natural disasters. The cultural heritage (CH) was identified as an important resource for building solutions to these scenarios in several international documents produced by the European Commission and the United Nations organizations (European Commission, 2015; UNISDR, 2015; Jigyasu et al., 2013; UN General Assembly, 2012).

CH is appreciated for conveying beauty and a sense of history, identity, and belonging for people, which contribute to the psychological well-being and quality of life of citizens. Additionally, the necessity is stated to promote an innovative use of CH for a smarter, inclusive, and sustainable civilization (European Commission, 2015).

The economic contribution given by CH is considered, obviously, as it concerns tourism, especially for developing a sustainable form of tourism, which can be reached by considering the several involved issues in a unique framework. Moreover, economic assets can be affected for attracting specialists studying CH (European Commission, 2015).

New society should exploit CH to foster inclusiveness and integration. The identity and belonging values of CH could be an opportunity for the engagement of people in society, and also an important issue for city resilience[10] to disasters, for the symbolic values they carry (Jigyasu et al., 2013).

As a new society is characterized by important cultural diversity, it is necessary to achieve a better understanding of local cultures and improved inter-culture dialogue for mutual understanding (European Commission, 2015). For this reason, a flexible representation is needed for adapting the documentation to specific situations and possibly unique cultural assets, and a common (standard) vocabulary and schema definitions are essential in order to reach a unique interpretation and effective understanding of reciprocal cultures in a harmonized framework. The archiving of CH documentation in a shared structure together with the reasoning possibilities of artificial intelligences could enlighten possible relations between cultural assets produced by different cultures and possibly identify some common background, thereby promoting inclusiveness.

A further role of cultural heritage, which is increasingly acknowledged, is the reference for environmental protection reasons: the traditional skills, practices, and know-how embedded in CH should be a base for a more sustainable development, building, and “building-back better” after disasters, further increasing resilience.

For these reasons, CH is also considered an important element in disaster prevention and management activities, as is explicit in the Sendai Framework for Disaster Risk Reduction 2015-2030 (UNISDR, 2015). To be effectively taken into account, it should be included in the tools analysing and managing the state of affairs, together with the other salient aspects and objects participating in determining risk, damages, and planning decisions. For this reason, it would be

[10] Resilience is ‘The ability of a system, community or society exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions’, United Nations Office for Disaster Risk Reduction (UNISDR), ‘2009 UNISDR Terminology on Disaster Risk Reduction’, Geneva, May 2009 (http://www.unisdr.org/we/inform/terminology).(UNISDR, 2015).

critical and strategic to document them using a common framework and common rules without losing the richness and uniqueness of each CH item.

The thesis proposal

In the light of these principles, this thesis deals with the management of the documentation of a particularly multifaceted branch of cultural heritage: the architectural heritage in both its spatial and thematic dimensions.

Standardisation is pursued in this research as an implicit requirement stated in the cited charters for fostering interoperability and an effective information retrieval.

The developing smart communities share similar goals. In this field, technologies are introduced as communication tools (by the internet community) and digital archiving methods. Furthermore, the field of artificial intelligences introduces algorithms and systems enabling new analysis possibilities and reasoning mechanisms that can improve the archived and shared knowledge with further inferenced information.

The current scenario for representation, storing, and sharing of information is well depicted in the words of the ISO/TC 211 strategic Business Plan about Geographic Information Standards (ISO, 2015). These are here reported as an effective synthesis.

The current environment encompassing geographic information and related standards is multifaceted. At the beginning, geographic information standards focused on the use and the integration of geographic information in any geographic information systems (GIS) no matter which software vendors that developed them. With the advent of the Internet and the Web[11]*, a massive amount of geographic information became readily accessible. Web services were needed for their discovery, visualization, access. Spatial data infrastructures (SDIs) emerged providing these Web services. Geographic information, traditionally acquired by governmental mapping agencies, is now also acquired by the population and distributed on the Web (e.g. OpenStreetMap initiative) as volunteered geographic information (VGI) or through crowdsourcing.*

Although geographic information, as well as non-spatial information sprung up independently over the Web, interoperability at the semantic level has been required to better understand them both by human and machine, and to perform additional reasoning (e.g., similarity assessment, inference, etc.). This has become possible with the Semantic Web which introduced the Web of data. In addition, the concept of Linked Data brought the capacity to connect and integrating together different piece of data, spatial and non-spatial, over the Web by the way of dereferenceable HTTP URIs breaking data silos and providing more information than before.

---

[11] The **Internet** is a massive network of networks, a networking infrastructure. It connects millions of computers together globally, forming a network in which any computer can communicate with any other computer as long as they are both connected to the Internet. Information that travels over the Internet does so via a variety of languages known as protocols.

The **Web**, is a way of accessing information over the medium of the Internet. It is an information-sharing model that is built on top of the Internet. The Web uses the HTTP protocol, only one of the languages spoken over the Internet, to transmit data. Web services, which use HTTP to allow applications to communicate in order to exchange business logic, use the Web to share information. The Web also utilizes browsers to access Web documents called Web pages that are linked to each other. Web documents also contain graphics, sounds, text and video.

The Web is just one of the ways that information can be disseminated over the Internet. The Internet, not the Web, is also used for email, which relies on SMTP, Usenet news groups, instant messaging and FTP. So the Web is just a portion of the Internet, albeit a large portion, but the two terms are not synonymous and should not be confused. (http://www.webopedia.com/DidYouKnow/Internet/Web_vs_Internet.asp)

Now Web sensors are providing real time information over the Web by the way of sensor web and Internet of Things (IoT) technologies. Earth Observation data is acquired continuously and is accessible on the Web. The new concept of Big Data and associated technologies that appeared recently brings new technologies to consume and analyze massive amount of data on the Internet which can be applied to the geo world.

Standards and geographic information standards are covering many aspects and domains of the current environment and resolve many issues, accordingly. For example, the World Wide Web Consortium (W3C) standards are covering much of the Web environment; the Object Modelling Group (OMG) is covering much of the data modelling aspect; the ISO/JTC1/SC32 is covering standardization in database including the spatial part. The Open Geospatial Consortium (OGC) and the ISO/TC 211 are covering standards specific to geographic information, such as metadata, encoding Web services, feature content, spatial and temporal schema, coordinate reference systems, imagery, coverage, gridded data, and so on and so forth.

However, many other issues have not yet been covered, for instance standards targeting domain specific areas and where standardized application schemas and ontologies need to be developed.[12]

The research of this thesis aims to respond to this last stated lack by filling the gap in the domain of architectural heritage representation.

The knowledge about architectural heritage is composed by two main categories: the thematic information and the spatial type.

Both can take advantage of worlds of already developed guidelines by parallel application fields. In particular, for archiving thematic information about cultural heritage stored in archives and museums, standards are published that are already used in several applications.

Cartography provides affirmed standards for managing spatial data. Moreover, Laurini (2014) affirms that for territorial intelligence and smart city planning, a new conceptual framework must be established by integrating artificial intelligence, knowledge engineering, computational geometry, and spatial reasoning. In his paper, he outlines this framework for geographic knowledge. Similar consideration must be granted for managing other kinds of spatial information. In the proposed thesis, the same principles and problems are studied for the architectural heritage knowledge management.

A method is proposed for archiving, analysing, and presenting architectural heritage data in an inclusive way (comprising multi-scale, multimedia, multi-temporal, multi-format information, connection with intangible values, technical analysis, hypothesis, and so on) by means of the use of ontologies. The already wide storage possibilities are improved by the introduction of the semantic values defined in existing standards (enhanced when necessary). This enables an unambiguous interpretation of the information and empowers the possibility of both database interoperability and information retrieval through the relations among the data defined in the data models. All the objectives stated in the Ename Charter can benefit from the advantages of these systems.

[12] The italic text cites (ISO, 2015)

ONTOLOGIES: INTEROPERABILITY AND AUTOMATED REASONING

A preliminary terminological explanation about the term 'ontology' is necessary in order to clarify to interdisciplinary communities, which is the sense intended in this thesis.

In particular, three definitions are possible (Guarino, 1998):

1) '**Ontology**', with the capital 'o' and in an uncountable sense, which refers to a philosophical discipline about the nature of being, existence, and reality;
2) 'ontology', with the lowercase 'o' and as a countable word, which can be in turn intended as:
    a. **ontology** in a *philosophical sense*, as a system of categories for a certain vision of the world; in this conception, it is independent from the used language, and it could be considered as a '**conceptualization**';
    b. **ontology** as the *engineering artefact* used in artificial intelligence (AI), constituted by a specific vocabulary and a set of explicit assumptions about the intended meaning of the words in the vocabulary. They are often called 'domain ontologies' or 'applied ontologies'.

In the thesis, with the term 'ontology', the last meaning is intended.

The domain ontologies (described in more detail in Chapter 5) are therefore an informatics tool that permits the definition of the semantics (that is the meaning) of the data. Applied ontologies build on philosophy, cognitive science, linguistics, and logic with the purpose of understanding, clarifying, making explicit, and communicating people's assumptions about the nature and structure of the world. They help people to understand each other, having an interdisciplinary nature. Slight differences in their definition can make them more suitable for one aim or another (coarse or fine-grained for establishing consensus, sharing representations, or making analysis, etc.) (Guarino, 1998). A long list of computer science activities that can be supported by the use of ontologies are presented, for example, in (Guarino, 1998). Among these, the knowledge engineering, knowledge representation, and information systems, especially geographic information systems, are the goals of this research. Moreover, the birth of the Semantic Web (from about 2006) brought new potentialities to the employment of ontologies in an effective way.

Briefly, their use permits the enrichment of the managed information, enabling semantic interoperability (realized as database interoperability and useful for information retrieval) and automated reasoning on the built knowledge bases (Laurini, 2015; Guarino, 1998). These are central to the thesis proposal.

In its most prevalent use in Artificial Intelligence (AI), an ontology refers to an engineering artefact (theoretical or computational) constituted by a specific vocabulary. It is used to describe a certain reality, plus the set of explicit assumptions in terms of primitive categories and relations describing the nature and structure of the domain of discourse (Guarino, Welty, 2000, 2002b). This set of assumptions usually has the form of a first-order logical theory[13], where vocabulary words appear as unary or binary predicate names, respectively called *concepts* and *relations* (Guarino, 1998).

Figure 1 is a schema that presents that definition of an ontology. To better understand it, the following definitions must be considered:

[13] https://en.wikipedia.org/wiki/First-order_logic

- a *conceptualization* **C** =<D, W, $\mathfrak{R}$>, that is, a set of conceptual relations ($\mathfrak{R}$) in the domain space <D,W>, where D is the considered domain and W is a set of states of affairs of such domains (possible worlds);
- an *ontological commitment* (intensional interpretation) for **L** (a chosen language, usually formal), that is **K** = <**C**, $\mathfrak{I}$>, where $\mathfrak{I}$: $V \rightarrow D \cup \mathfrak{R}$ is a function assigning elements of the domain D to constant symbols of a vocabulary **V**, and elements of relations $\mathfrak{R}$ to predicate symbols of V (in other words, it is a conceptualization expressed by means of a language);
- the whole of the *models* **M** considered as possible interpreted world structures, in a defined language **L**;
- the subset of the *intended models* **I**, that is, the models of L according to K;
- the *ontology*, which, for a language **L,** *approximates* a conceptualization **C**, when an ontological commitment **K** = <**C**, $\mathfrak{I}$> exists such that the intended models **I** are included in the models of **O**.

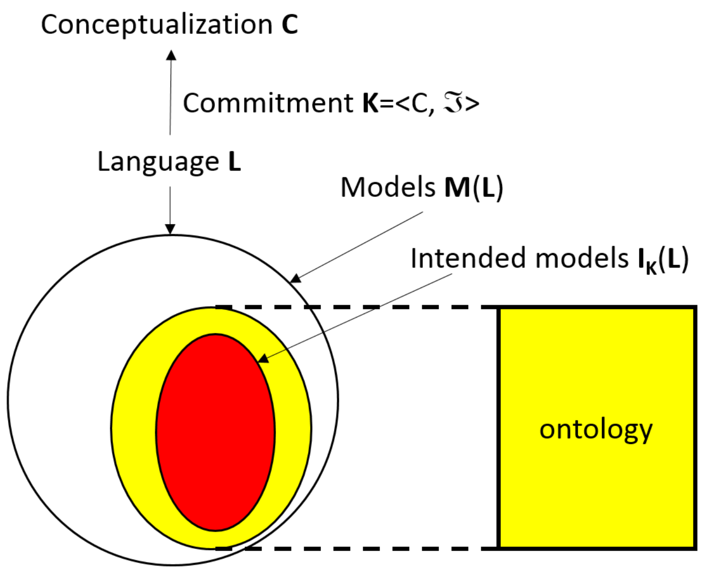


Figure 1. Schema illustrating the relations between an ontology and the aspects to be considered in a representation or description of a part of reality. (Guarino, 1998).

In this mechanism, the choice of a suitable domain and vocabulary must be considered: the degree to which an ontology specifies a conceptualization depends on the following factors (Guarino et al., 2009):

- the richness of the domain of discourse (for the architectural heritage, this richness is clear);
- the richness of the chosen vocabulary (Figure 65 depicts the expressiveness of language);
- the axiomatization (the reduction of the real-world systems to ontological axioms).

In order to be understandable and reusable, the ontologies must combine the precision of formal semantics with the efficacy of cognitive transparency (Guarino, Musen, 2005).

Reasoning and machine learning

AI disciplines, which employ ontologies, developed several techniques for inferring further information from certain input data.

Some cognition and processing capabilities are here described, which are generally typical of the human reasoning, and can be implemented in some automatic tools, such as Geographical Information System (GIS) software tools (Worboys, Duckham, 2004).

One of these is reasoning: the process by which existing information is used to infer new information about a problem domain.

1) The '**deductive**' inference (applying rules to specific examples), which is characteristic of the typical case of Aristotle's syllogism[14]

[rule + case => result]

(ex. All British people like tea, Mary and Mark are British, so Mary and Mark like tea).

2) The '**inductive**' inference (generalizing from specific examples to general rules)

[case + result => rule]

(ex. Mary and Mark are British, Mary and Mark like tea, so, all British people like tea).

3) The '**abductive**' inference (generating explanations for some state of affairs)

[rule + result => case]

(ex. All British people like tea, Mary and Mark like tea, so, Mary and Mark are British).

The rules of deduction preserve truth: if the premises are true (the knowledge base is 'sound'), the conclusion must be true. A knowledge base is complete if all true propositions are provable by deduction from its premises.

Inductive and abductive inference are central to human reasoning, despite their unreliability. They allow people to reach conclusions and generate further hypotheses. Some artificial intelligence-based systems use induction and abduction to extract new information from data. Machine learning techniques, for example, use this sort of reasoning. As humans do, machines could use some mechanism for regulating unreliable reasoning, for example by incorporating implicit contextual knowledge in the form of constraints.

Machine learning is a part of computer science that deals with algorithms that can learn from wide datasets, by calculating statistics and making associations in the known datasets, and associating the similar characteristics in new data to the learnt concept.

For example, machine learning[15] techniques can be effective for this aim, because they permit not being limited to deductive inferences but have the advantage of some more complex reasoning. Machine learning techniques can individuate characteristics in large datasets, and abstract them as rules for interpreting further similar information in other input data.

[14] http://plato.stanford.edu/entries/aristotle-logic/#AriDedModValArg

[15] https://en.wikipedia.org/wiki/Machine_learning
http://whatis.techtarget.com/definition/machine-learning

Another field that can be improved by the introduction of artificial intelligence in database manipulation is problem-solving. Some reasoning concepts should be used, such as the heuristics, the analogy or metaphor, and learning (Worboys, Duckham, 2004).

SECTION 1 – THE RESEARCH OBJECTIVES

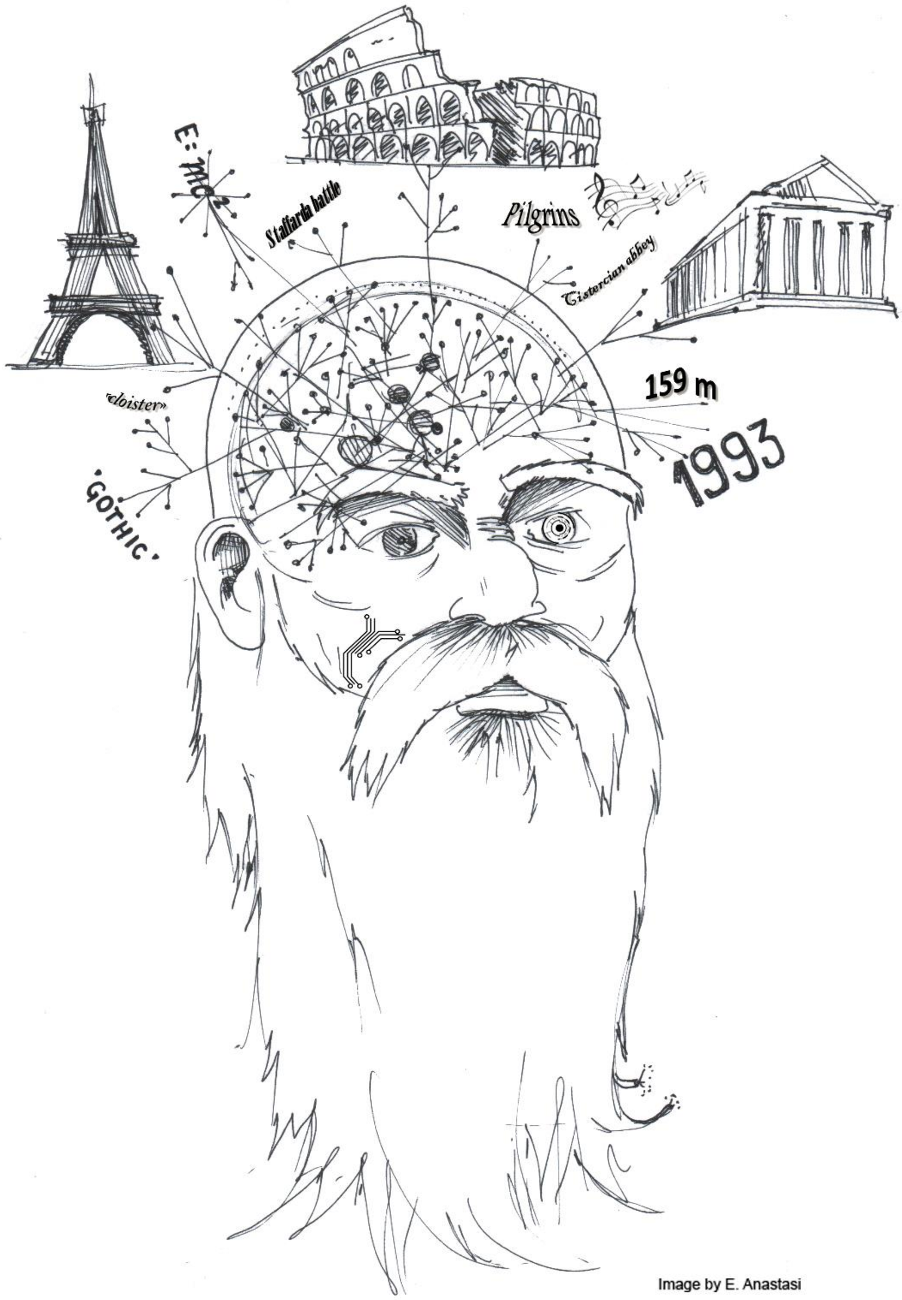


Image by E. Anastasi

1 THE CHARACTERISTICS OF ARCHITECTURAL HERITAGE DOCUMENTATION

As widely recognized and stated in the world of cultural heritage and architecture, and also perceived by each common citizen, architectural heritage has a number of aspects to be considered for being effectively documented and represented. It has multifaceted characteristics and a complex nature, which links monuments to a wide range of further themes and concepts. First of all, there are thematic aspects that must be considered, since they are often the reason of the 'Cultural Item' statute (i.e. they represent significant cultural issues strictly connected to the object). In the case of architecture, they acquire consistency if they are related to the material manifestation of the artefact, whose metric survey and documentation is unavoidable for correct analysis and representations.

The needs for recording and documentation of cultural heritage are summarized in (Stylianidis et al., 2011). The data have three dimensions plus the time reference. They must be recorded following the necessity of their nature, which is multi-format (from textual data to images, historical photographs and 3D models), multi-content, and multi-source (more than one source can document information in similar or dissimilar ways, but it is important not to exclude none of them). The data should be managed in digital inventories to be shared and distributed to the users.

The quality of the managed information is critical for obtaining meaningful results. This should be considered from a double point of view: the goodness of the measurement or acquisition and, on the other hand, the semantic quality of the representation. For both it is possible to define a precision and accuracy/correctness to be considered in the documentation (Figure 2). Similar concepts are stated in the field of measurement quality and reference ontology quality. Accuracy and precision are schematized in Figure 2.I, where the red point is the point to be measured, and the yellow points are the set of performed measurements that approximate the correct one using an instrument in a measurement operation. Semantic precision and correctness are instead represented in Figure 2.II. Figure 1 explains the meaning of the schema parts (the whole of possible models in grey, the ontology in yellow and the intended model in red).

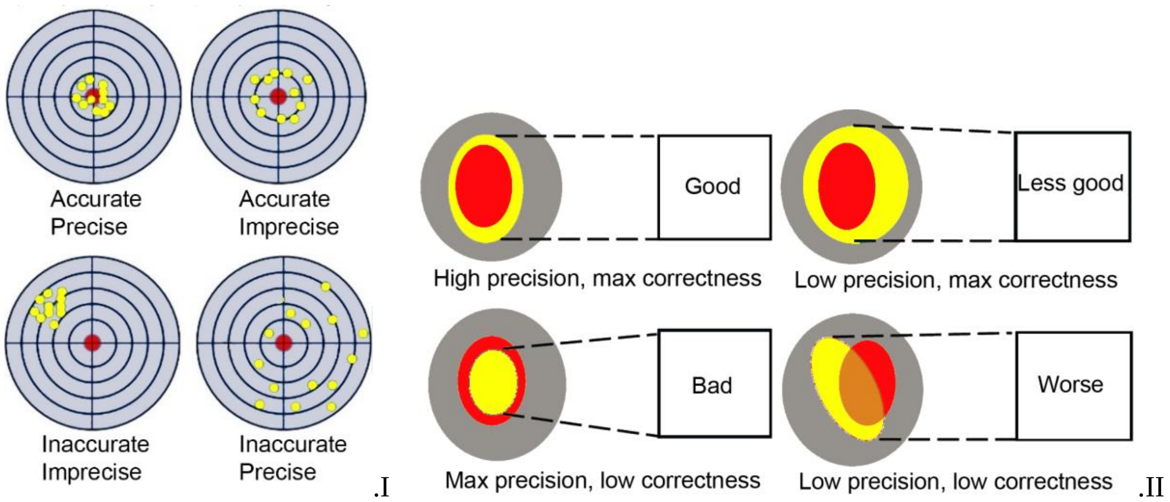


*Figure 2. Schemas of the similar concepts about representation accuracy in the field of **measurement quality** (I) and **reference ontology quality** (II). A similar colour-code is used in both sets of diagrams: in red is the representation objective and in yellow the representations.*

Another important aspect is the possibility of the effective visualization and presentation of the collected information in a user-friendly way, possibly via the web.

Some challenging opportunities in cultural heritage documentation can arise from the data fusion from heterogeneous sources in a unique framework and the introduction of some standardisation issues (Stylianidis et al., 2011), so that the complexity cannot be lost, but instead emphasized to become a strength.

1.1 The thematic side

Many disciplines are involved in the study of cultural heritage and architectural heritage.

An architecture is first of all an artefact, of which it is important to document the shape, the materials, the appearance, the composition of its parts, the used building techniques, the spaces arrangement, and all of what is connected to the material object itself. This includes the history of the building: construction, changes, restoration interventions, demolitions of all or part of the building, and persisting pathologies (with connected causes and possible solutions).

A wide set of concepts is also connected, including the classification of the building and of all its parts, the decoration elements, the historical events connected to the building, the authors and actors in the building process, the meanings (from a collective or individual point of view) linked to the building or its parts, the interpretations about style, urban connections, intangible values and heritage possibly connected to the building, etc. Moreover, for architectures, the attributes connected to the present functions and role of the building are to be considered (present use, present functions, held events, economic activities, economic management data, promotion activities, energetic characteristics, visiting itineraries, safety issues, connected urban dynamics, and so on).

The documentation of this information can be effectively managed in relation to the metric products to which the data have to be referred for a powerful interpretation and reading. In describing these characteristics, some additional factors must be considered that are fundamental for cultural heritage documentation. One aspect is the fuzziness of the data (some data can derive from one or more descriptions or documents, but the degree of reliability of the sources and their interpretation can be various). A second fundamental aspect is the time. The data must include the indication of the time they are referred to in order to be meaningful for a correct reading of the information and also for running analysis, modelling and monitoring of the artefact.

Moreover, it is important to know the main function, the possible alternative uses, the economic values, the owners, the authorities having some influence on it, and the logistic aspects that can be the source of promotion activities and could aid the preservation of the monument. As stated by the RecorDIM documents, understanding a building or site '*involves understanding a number of social, political, economic, and cultural issues relating to the external environment*.' For this reason, a multi-disciplinary team of experts is always necessary.

A further thematic information source can be considered certain data, that give a visual description of the object, even if they also have a metric value, such as the photographic texture applied to 3D models, which permits the experts to understand and individuate some visual feature. Also raster products generally produced by GIS spatial analysis about the characteristics of a surface can be considered as thematic information (they can, for example, describe the irregularity of a wall due to static problems). Another possibility, which is being developed, is to use the multispectral acquisitions (thermal images included) to survey further information, which is useful for different aims in the study of artefacts (Hemmeleb et al., 2005).

For documenting the thematic values of cultural heritage (generally without including the last listed ones, more linked to geomatic disciplines), national and international institutions have developed cataloguing systems and schemes that are detailed in the sections addressing the standards about cultural heritage, later in the text (Sections 6.2 and 7.8.1).

1.2 The geometric side: from cartography to architectural large scale

In the years 1999-2000, architecture preservation operators and researchers proposed a document as a base towards the ‘Charter of the architectural survey’ (Almagro et al. 2000). In this, some criteria for the planning of the survey, the integration of the techniques, the management of synthetic and interpreted data in databases, and operational archives for architectural heritage are proposed. Here, the possibility to manage the data using information systems is encouraged.

It should be noted that the development of theories, techniques, and methods for the survey and management of geometric data comes from geography and geometric disciplines.

The metric aspect in the architectural heritage documentation is fundamental: the study of the architecture passes through the knowledge of the shape of the building. The relationship between the parts, the proportions in the parts and the spaces, the dimensions in comparison with the human being and the urban or landscape conformation are meaningful characteristics for the definition of an architecture and its cultural connotation. These aspects can also serve as the base for establishing cross references among different buildings and architectural works, bringing to light some reasons or some meanings, that help in understanding the correct interpretation of the artefact and the implied human or urban dynamics.

The recording of the shape of a building permits the identification of the more abstract lines from which the project and following stratifications originated and, on the other hand, to verify the used building techniques. Nevertheless, the recording of the real present geometry of the building is essential in order to survey the building pathologies of the materials and of static components.

Some further characteristics of the geographic information are finally considered (Laurini, 2014), which can be valid also for many aspect of the spatial management of the monuments. The awareness of these aspects is preliminary to effectively building spatial knowledge bases. The list of the most meaningful ones to our aims follows.

- *Link of geographic semantics and scale*. Different visions are useful for different aims; therefore, different features must be represented.
- *Multiple representation*. Different kinds of representation of an object can be required: a building can be represented as a point, a 2D polygon, a 3D solid or a 3D set of surfaces, based on requirements.
- *Multiscale*. Similarly to the previous point, different geometries can represent the same object, depending on the scale. For example, a 3D mesh can have a varying number of triangles depending on the representation scale, the computational capabilities of the hardware, and the necessity of detail.
- *Difference in classifying real features*, e.g. ontologies.
- Toponymy and *terminology* in parts definition. The relation between the object and its name is many-to-many, due to the different languages, the time, and some cases of homonymy.

This can be a thematic issue, but since the problem is already treated for geographical names, it is cited in this section.

- Necessity of interoperability among systems.

In the same paper (Laurini, 2014), the fuzzy geometric objects are cited as particular geometric objects. From a cartographic point of view, a fuzzy object may be an area where sometimes a represented entity or phenomenon exists and sometimes not (the example is a riverbed). One can therefore represent points or areas that belong to an entity with a certain membership grade (Laurini, Thompson, 1992). For architectural heritage representation, the fuzziness of information involves different aspects. For example in the reconstruction of parts of a building from historical sources, different geometries can be considered, depending on the different literary or drawing sources and on the persistency on the site. In other cases, for example a deterioration mapping can have some fuzzy areas in a conception similar to the previously described geographic one. In any event, the fuzziness of an instance of information should be included in the knowledge base. This data can be considered when querying the system and running inferences with the inclusion of the fuzzy logic theories, and can be represented in the visualisation, for example with different degrees of transparency or similar tricks. More fuzziness (both for thematic and spatial aspects) can derive from the processes of the source of information integration possible in the web. Some studies about the integration and uncertainty visualization of the resulting information are presented in (Berjawi et al., 2014).

A further aspect to be considered is also depicted in (De Luca et al., 2008), which defines the architecture shape to have two natures: the ideal and perfect design of the project and the concrete and imperfect real realization. The architectural survey can be equally oriented to the two goals. The first is the retrieving of the design model of the real building, for which a Building Information Model (BIM) approach, like the one proposed in (De Luca et al., 2008), could be perhaps more effective. Another one is the documentation of the reality in an imperfect digital model from which defects, pathologies, and degradation could be represented and highlighted.

Architectural survey and cartography have been considered for a long time as two distinct disciplines, but as enlightened by the theories about spatial knowledge, and some additional considerations, the two domains can effectively have advantages from a reciprocal influence (Hurni, Sell, 2009; Guerra, Miniutti, 2000). Geographic standards, methods and tools can be used for improving architectural surveys and 3D modelling. In the meantime, survey and archiving techniques of small objects (in relation to cartographic scales) can improve cartographic quality and content.

1.3 Further elements for spatial knowledge

The previously described aspects mainly regard geometry, which requires measurements and coordinates.

However, spatial knowledge is not only characterised by the metric aspect. The relation between the parts is also important to be considered. Some non-geometric data can be useful to be represented for permitting the relation of the object to other representations or to other objects. (De Luca et al., 2008) states the architecture representation to be composite by a geometrical description level, a topological relation level, and a spatial relation level.

Moreover, additionally for the different levels of definition of some objects or phenomena, some fuzzy logic concepts in the representation should be considered.

These topics concur to the constitution of a geographic knowledge (Laurini, 2014), which is in turn part of the architectural heritage knowledge itself. Nevertheless, the representation of architectural 3D models and their parts should consider very similar issues.

The conceptual framework for geographic knowledge is analysed by (Laurini, 2014). In that paper some prolegomena (preliminary assertions constituting the underlying foundation of principles) and principles are identified, which concur to the passage from a simple geographic representation (including geometrical aspects and some more non-metric information) to the building of a more complex geographic knowledge. The passage brings the advantage to provide tools for spatial database interoperability, reasoning and geographic information internet retrieval. The elements of knowledge are asserted to consist of the following: facts (data or instance); concepts (classes of items); processes (flows of events); and rules (to make inferences or draw implications).

The same issues should be considered twice for architectural heritage representation: once for the geographic characteristics and location of the monument, and once for the representation of its own parts considering a larger representation scale.

Geographic objects are characterized by having mutual relations. Some of them are typical of any knowledge base, such as the relations 'is-a' (hierarchy), 'has-a', 'part-whole' (Figure 7) (mereologic relation among the parts and the whole that contains them). To these ones, some more should be added (Laurini, 2014). In particular, for large-scale representations, topological relations should be considered, plus particular mereological relations among non-connected parts (for example, among the islands and the State they belong to, or the parts of buildings and the whole complex). Topology and mereology will be considered in detail in the next specific subsections (1.3.1 and 1.3.2).

In the following, a meaningful selection of these prolegomena and principles are reported (Laurini, 2014) considering their importance also for a representation of the monument itself.

Prolegomenon 1) 3D+T objects: All existing objects are tridimensional and have temporal evolution. The possibility to maintain these dimensions in the representation is of primary importance for monument representation.

Prolegomenon 5) (From Popper's falsifiability principle) *When a new apparatus delivers measures with higher accuracy, these supersede the previous ones.* For architectural heritage this is valid if some errors were observed in previous survey or some other cases. But it is often necessary to keep the previously acquired information stored in order to maintain historic data that permit monitoring consideration.

Prolegomenon 6) Updating should be done permanently.

Prolegomenon 7) All geographic databases or repositories must be accompanied with metadata.

Prolegomenon 9) A good practice should be to store all geographic objects with the highest possible accuracy and to generate other shapes by means of generalization. This preliminary consideration is often valid when realizing a building survey with the described techniques (laser scanning and image processing modelling), since their efficiency permits the storage of high quantities of data in reduced times; therefore, it is economically advantageous to acquire superabundant data, to be *a posteriori* reduced, filtered, and classified. The representation scale that is often considered as objective when realizing a survey is approximately 1:50. However, some care must be taken in the maintenance of some definition not to be lost in the approximation passages (e.g. the principal edges or the general shape).

Prolegomenon 10) Relationships between places and place names are many-to-many. This was already cited in the introductory section, and can also be valid also for the terms identifying some specific parts of the object.

For solving this problem, the reference to gazetteers are used, in order to disambiguate the relation between the place and its name. They are directories of names (Laurini, 2015), in particular toponyms, which can also include more complex information[16]. Their scope is to collect the set of names given to a place (or, translating the concept, to an object), for solving multilingualism ambiguities, multiplicity of similar terms, the changing of the names over the time, and homonymies. They can be simple lists of equivalent terms or names, unambiguously related to an ID, or they can contain more complex information (see Laurini, 2015 for some examples).

Some examples of gazetteers or similar systems are GeoNames (Section 7.8.4), and the Getty Thesaurus of Geographic Names (Section 7.8.2.3.2) for geographic toponyms, while, for example, the others Getty vocabularies[17] (Section 7.8.2.3) can be considered as reference for the definition of further art and architectural terms or related issues.

Some inference rules can be applied to geographic knowledge bases for matching geographical ontologies in different languages by considering the homology relations comparing their geometric shape or equivalence of toponyms (Laurini, 2015). Furthermore, some markup languages were built, (such as SpatialML) to geo-code the natural language texts (Mani et al., 2008). This implies that they could be related to places or 3D models when the entire system was operating. A gazetteer service is also included in OGC web services guidelines and OGC best practises (OGC, 2012b).

Prolegomenon 11) All geographic object types are linked to concepts organized into a geographic ontology based on topological relations. The same is true for the parts of buildings and monuments. This prolegomenon (described in greater detail in Laurini, 2012) underlines the issue that is the subject of this thesis. Further definitions of ontology will be given in Chapter 5, and existing data models to be considered as common ontologies will be treated in the research parts.

Prolegomenon 12) Everything is related to everything else, but near things are more related than distant things (Tobler's law). This law should be further studied for how it concerns cultural heritage. Undoubtedly, the common analysis methods in studying monuments consider similar cultural heritage instances in neighbouring areas before considering relationships to far places. However, the possibility to analyse similar artefacts in remote locations should underline further knowledge elements (Noardo, Spanò, 2015). This is also due to the complexity of social and anthropological aspects that are intrinsically linked to cultural heritage. A global knowledge base would permit the better investigation of this topic.

[16] https://en.wikipedia.org/wiki/Gazetteer

[17] http://www.getty.edu/research/tools/vocabularies/

After these preliminary considerations, a set of principles governing geographic knowledge are stated (Laurini, 2014). Again, some of them are selected and considered also from the representation of thearchitectural artefacts point of view.

Principle 1) Spatial knowledge is hidden in geometry whereas geographic knowledge comes in addition from non-spatial attributes. As stated in the introductive part, the knowledge about architectural heritage is composed by both thematic and spatial information.

Principle 5) Cartographic representation is linked to visual acuity. This opens a study issue for what concerns the simplification of a 3D model in function of the visualisation requirements and computational capabilities.

Principle 7) Spatial relation varies according to scale. For example, what is in touch or what is disjoint varies in relation of the representation scale. Considering a 3D representation, these aspects are possibly more complex, especially if the reduced model comes from a denser one.

In the next subsections, two particular and widely investigated issues important to spatial data management are explored: topology and mereology.

1.3.1 Topology

Topology is the mathematical study of the properties preserved through deformations of objects. Some concepts, in particular topological relationships, are important in the GIS domain (Brahim et al., 2015).

(Allen, 1991) proposes a model for the topological relations between segments, which can also be considered as a metaphor for temporal representation.

(Egenhofer, Herring, 1990) proposed a mathematical framework for dealing with topological relationships based on geometrical primitives. They proposed to consider the possible combinations of six topological primitives (considering the interior, exterior, and boundaries for two 2D objects). The results of all the possible combination among the six primitives are the 8 topological relations shown in Figure 3.

Further studies were proposed by (Lee and Hsu, 1992) for the definition of topological relationships between rectangles (Brahim et al., 2015).

Other studies define theories for fuzzy topology (Solovyov, 2012), which could be of interest for cultural heritage representation, as well.

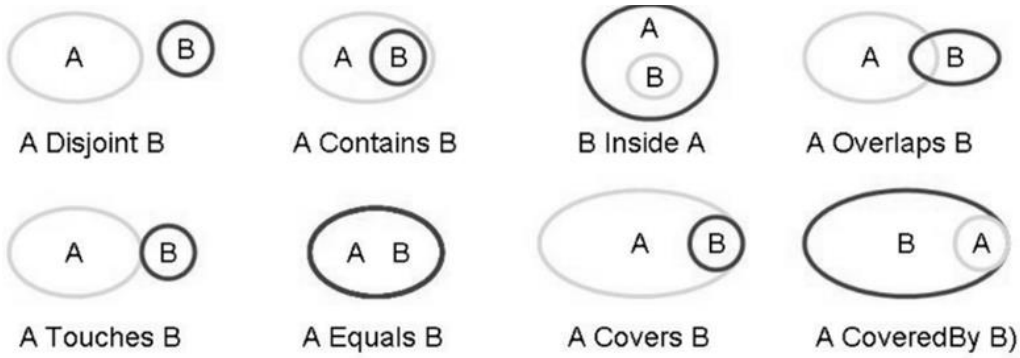


Figure 3. Egenhofer topological relations. These should be integrated into ontologies' definitions of geographic features (Laurini, 2015).

Starting from similar concepts and applying some theories for the relationship verification, some more topological relationships are defined for 3D objects and 3D space (Zlatanova, 2000). Other investigations aim at the application in 3D GIS (Ellul, Haklay, 2006; Zlatanova et al., 2004; Stoter, Van Oosterom, 2002) and 3D objects in general (Leopold et al., 2014; Hmida et al., 2013) for enabling advanced spatial reasoning. Some of them are represented in Figure 4 and Figure 5. Some of these concepts have been applied in some cases for 3D building information management (Billen, Zlatanova, 2003). Moreover, some studies focus on the inclusion of topology in semantic 3D city models (Bucher et al., 2012). The methods proposed by these investigations should be tested in future work structuring the analysed 3D models by the enhanced data schemas developed in this thesis.

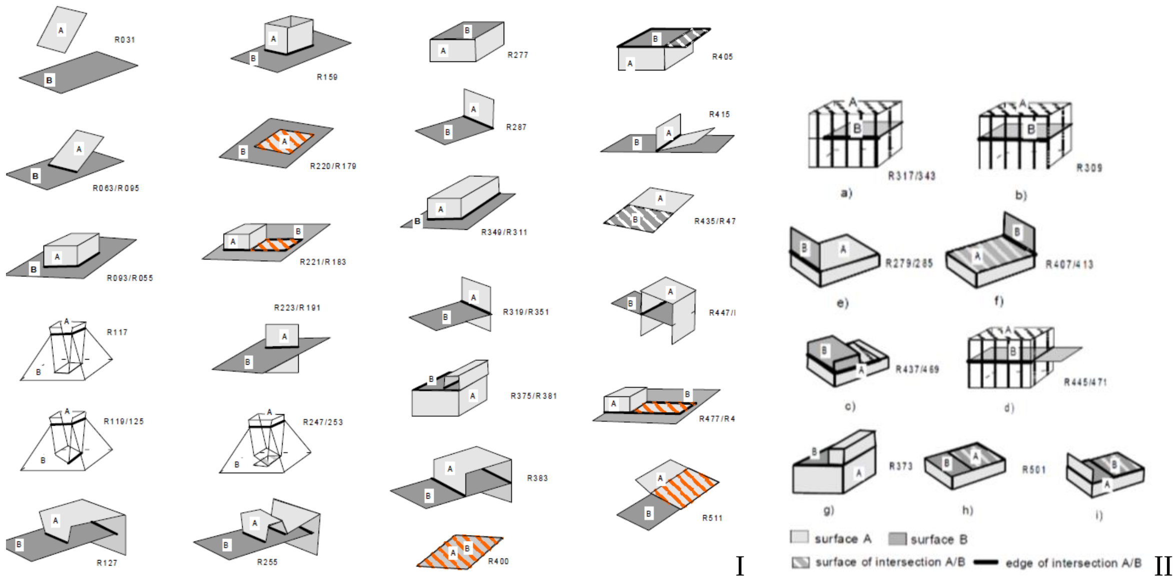


Figure 4. (I) Some of the possible surface-and-surface relationships. (II) Examples of closed surfaces containing a volume (Zlatanova, 2000)

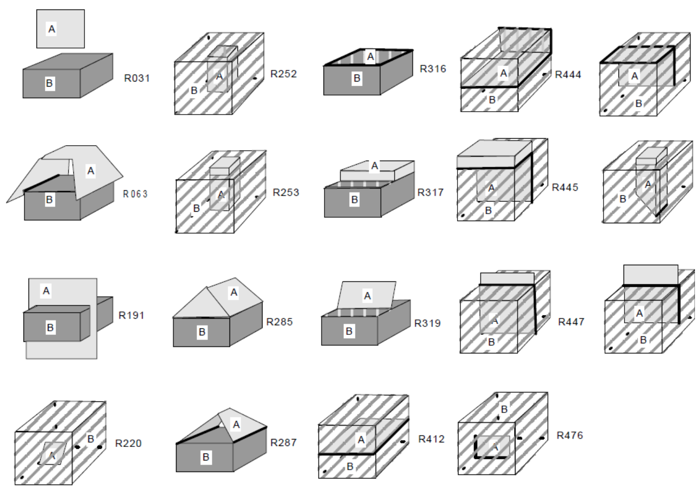


Figure 5. Body and surface in IR^3 (the 3D space): 19 relationships possible for a body (Zlatanova, 2000)

1.3.2 Mereology

Mereology is the branch of ontological investigation dealing with the part-whole relation (Keet, 2006). Mereological studies give some interesting considerations about the part–whole relationships, which are essential to spatial data management. In addition, a logic and ontological definition is given.

It is important to consider mereology in this preliminary chapter since the exposed concepts are meaningful for a correct documentation, even if the given definitions are mainly object of study of specific fields, as a branch of logic and ontology, treated in other sections (5.2). Its themes are useful to conceptual modelling, since partonomies contribute to characterize semantics and provide mechanisms for automated reasoning (e.g. if a head, a body, two arms, and two legs are part of something, this something can be a man; if a nartex, a cell, a pronaos are part of something, this something can be a temple) (abductive reasoning).

Mereological relations are classified in different sub-relations, following a sort of hierarchy, shown in Figure 6. Moreover, a distinction must be made from the meronymic part-whole relation, which is not transitive (ex. A grain of sand is part of clay, clay constitutes a vase, but a vase is not constituted by grains of sand).

Transitivity is the property for which a property can be transferred to subclasses; for example, if A is part of B and B is part of C, A is also part of C.

The *part-whole* relations are classified as:

A. **Mereological** relations:

1) ***parthood***: it determines the general part-whole structure of an object (ex. palace – left part of the palace). It is dense, due to the existence of 'fiat' parts (parts that lack a complete bona fide, that is well-defined, boundary); this is called also 'portion-object' by Odell (Keet, 2006) and in the same reference, the 'place-area' relation can be considered as very similar (***portion – object***, this can be some amount of matter is part of the whole, or scale-based partonomic relations; ***place – area***, where part-place cannot be separated from the whole area);

2) ***componenthood***: a component is a proper part of that object that has a 'bona fide' boundary (a boundary that corresponds to discontinuity in reality) and a distinct function (ex. column – capital). It is discrete: if *x* is contained-in *y* then either:

- *x* is immediately contained in *y*
  or
- there exists a *z* such that *x* is immediately contained in *z* and *z* is contained in *y*
  or
- there exists a *z* such that *x* is contained in *z* and *z* is immediately contained in *y*.

The same relation is called by Guizzardi (2011) ***component – functional complex*** (engine – car) and by Odell in (Keet, 2006) ***component – integral object***, that is, a discrete type of part-whole relation with atoms;

3) ***containment***: relation between disjoint material objects when one object is located within a space partly or wholly enclosed by the container (ex. car, driving area; church, sacred area). It is discrete, as well;

4) ***subcollective – collective*** (underage children of John – John's children; Savoy Residences – Royal Residences) (Guizzardi, 2011). This is not expressly identified in Figure 6, but it is considered a mereological relation, since it is transitive.

B. **Meronymic** relations (not transitive):

5) ***material – object,*** constitution of objects (Odell, in Keet, 2006);

6) ***subquantity – quantity*** (ex. alcohol – wine) (Guizzardi, 2011) (further studies on this particular relation are made by Guizzardi, 2010);

7) ***member – bunch***, an entire bunch is generally denoted with a collective noun and its members can change over time (Odell, in Keet, 2006); it is also called ***member – collective*** (ex. tree – forest) (Guizzardi, 2011);

8) ***member – partnership***, like member – bunch, but changing a member does destroy the whole (Odell, in Keet, 2006).

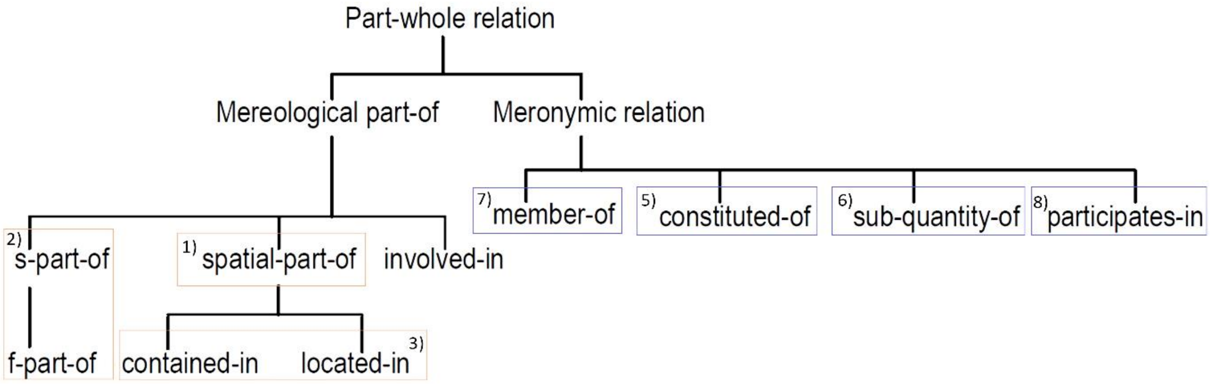


Figure 6. Taxonomy (hierarchy of terms) of basic mereological and meronymic 'part-of' relations. 's-part-of' stands for 'structural part of' and 'f-part-of' means 'functional part of' (Keet, 2006). The numbers refers to the description in the numbered list above. The relation 'involved-in' relates to parts of a process, which are not considered here.

Mereology theories are fundamental for the study of architecture composition and the meaning this carries (Figure 7).

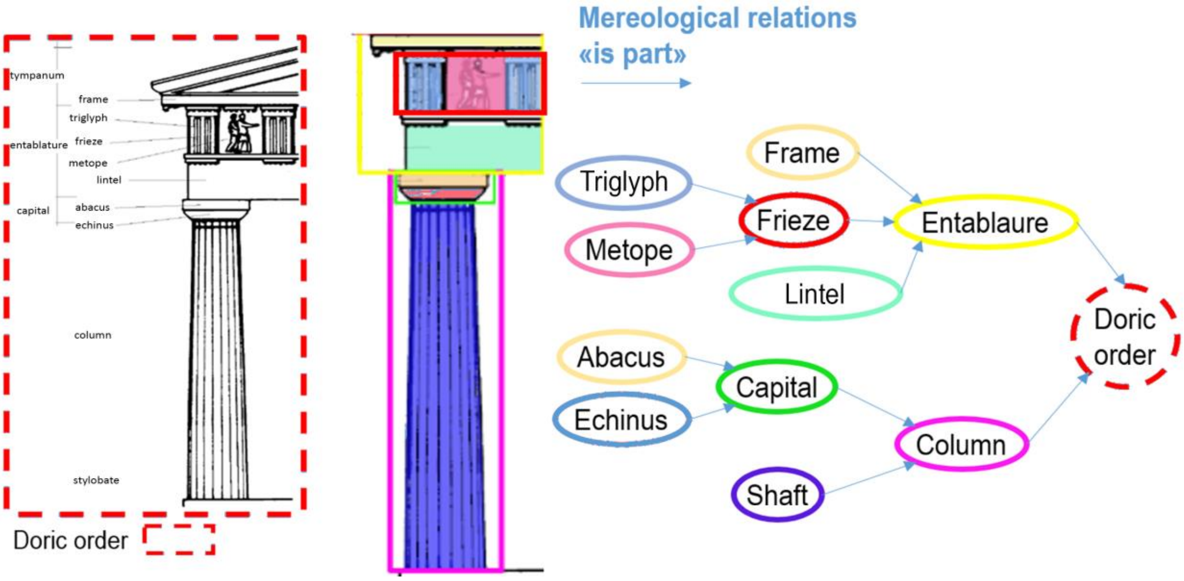


Figure 7. Example of 'part-whole' relation on a doric order.

Mereological hierarchies of architecture parts are investigated especially in the archaeology field, where the documentation of small pieces and details is fundamental. Also, in the study of architecture the hierarchical interpretation of building elements is an affirmed practise. However, its formalization for the management of digital data by means of informatics applications was little considered. Some innovative studies in this sense were performed, for archaeology first, in order to move in the direction of a representation and information sharing through web archives. Some investigations about the semantic segmentation and storing of digital data for representing and archiving the documentation about archaeologic case studies have been published (Apollonio et al., 2010; De Luca et al., 2008; De Luca et al., 2007). Some of them use CAD-based enhanced software, such as Autodesk Maya (Apollonio, 2010). In other cases, specific software applications were programmed for managing such data including multimedia materials. One semantic-based web platform that could be used to publish semantically annotated data was 'Nubes'. It implemented hierarchies to be used for annotating the inserted data (De Luca et al., 2011; De Luca et al., 2011b). A similar implementation using open source software was performed by (Guarnieri et al., 2010).

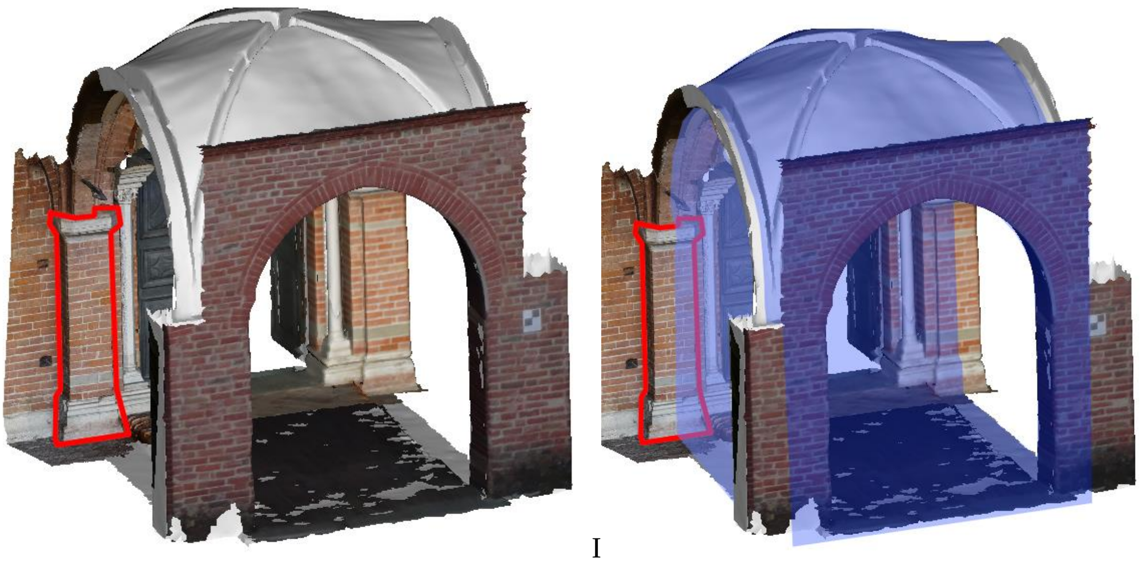
The issue is also studied in cartography, within which the coverage of one area by subareas is called "tessellation"; it can be hierarchical (as in the case of administrative units). Translating the same concepts to the 3D representation of a building, some consideration should be done.

First of all, the different levels of detail and consequent shape simplification produce surfaces that do not coincide perfectly, even if representing the same object (Figure 8). Therefore, some intermediate 3D space should be used for performing spatial queries; for example, the envelope of the object, or a volume at a certain approximated buffer from the surface, including the surface in the middle.

.I .II

Figure 8. Different levels of detail of the same object, lower in (I) and higher in (II).

A second consideration is the existence of some partition elements that have their own identity at a lower hierarchical level and should be considered to be split in a superior hierarchical level (see the example in Figure 9 and Figure 10).



I II

Figure 9. In an inferior hierarchical level a pillar has to be considered as a unit (I); but it should be split into two parts when considering a bay (blue in II). The fact that the two objects could be considered for two different classifications (a constructive subdivision for the pillar and a spatial composition approach for the bay) could be objected, but anyway it is an issue that should be discussed.

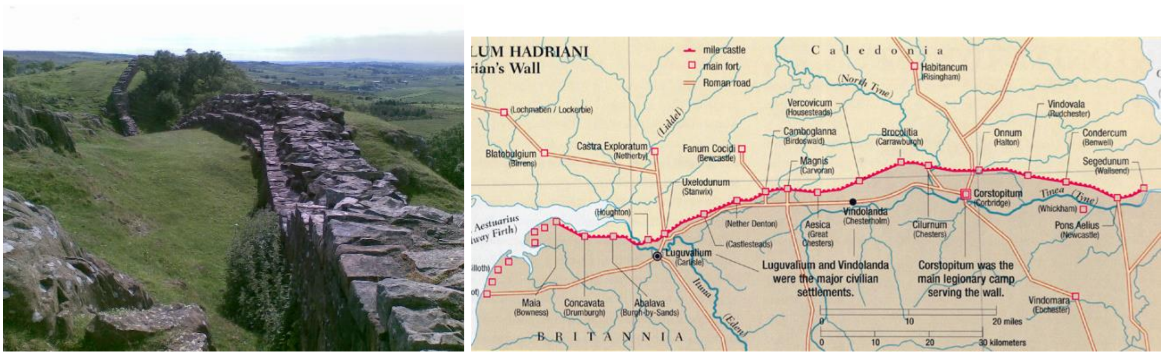


Figure 10. Lower-scale example: where did Britannia end in a 3D representation of Hadrian's Wall? This should be defined or decided. (The images were retrieved from the web[18]).

1.3.3 Geo-spatial rules

In recent papers (Laurini, 2015b), the problem was presented to store more geospatial business rules, which should consider further knowledge (social, economic, political, etc.) than simple spatial ones, for helping in geographic reasoning. Examples of this are: 'mosques are oriented towards Mecca'; 'the prices of houses in more touristic places are higher'; 'the antique towns were mainly located on high places or near rivers, but with exceptions', and so on. For monuments, similar rules should be stored, for example, for what concerns the Catholic Church precepts on churches' architecture and symbology in various epochs, or for accessibility to various building spaces for men, women, etc.

[18] http://pliberat.altervista.org/parent/inghilterra/112.jpg and http://www.progettolimes.it/wp-content/uploads/2014/03/Map-of-Hadrians-Wall.jpg

Some of these business rules could be stated using the `if-then` syntax. Another method can be the use of Object Constraint Language (OCL), useful also because it is expressly proposed as an extension of UML, which is the language used by most geographic data models. Some study of this topic, also for validation purposes was made by Wagner et al. (2013).

In (Laurini, Thompson, 1992), the importance is stated for the spatial data representation to embody not only coordinates and topology, but also geometric semantics such as rules, since algorithms and methods are used in data modelling as rules.

1.4 Time

The time representation management should be the object of a specific investigation, since a number of topics should be considered. Some of them for historical data are: the methods for determining the time to be associated to the data, the validation issues, the wideness of the period to be considered, the time reference system to be used, the various degrees of fuzziness in the determination, the topological relation between periods, and so on (Kauppinen et al., 2010).

On the other hand, acquisition sensors and techniques are evolving towards almost real-time acquisition – processing and communication processes. Consequently, large dynamic series of data are acquired and stored for monitoring aims. For this reason, specific standards and protocols are being developed in representation and management fields for taking advantage of this great potentiality.

In both cases a model for structuring information, and specific formats and methods should be separately treated.

1.5 An example: an architectural part and reasoning mechanisms

It would be of interest to store the presented components of spatial information in 3D models of monuments for enabling automatic inferences or machine learning potentialities on architecture parts, from some trivial ones (like the one presented in this section) to more complex information extraction. Very interesting research for the matching between 3D models of semantically coincident objects basing on topology is described in (Hilaga, et al., 2001). The system here presented would offer the potentiality to exploit actual tools for performing similar matching on world-wide scales, once the data are structured and published on the Semantic Web.

An example of spatial information defining an architectural object (a gemel window) is here presented (Figure 11). The thematic and geometric information is in this example not considered, for illustrating how topological, mereological and spatial information could be considered for inferring unknown information.

- The **topological information** could be defined (in natural language) as:

  ‘it contains two similar windows (with similar dimensions and proportions); the two windows touch.’

Alternatively, at a larger scale (considering the principle 7 in Section 1.3):

'a window touches a vertical element, which in turn touches another window on the other side.'

- Other definitions can regard the **shape** (the interior topology) of a window:

'it approximates a rectangle (or parallelepiped in 3D), generally disposed vertically, considering the main direction, and whose approximated proportions can be indicative of the epoch of construction or used building techniques.'

The definition of a window, since it is a void, is another issue, but this will be treated by studies involving segmentation techniques.

- **Mereological relations** are:

'the two windows and the vertical elements and, eventually, the filling element on the top, etc., are part of the gemel window.'

All these descriptions are common to all the gemel windows; therefore, they can be used to make inferences and to accomplish information retrieval in a database containing several similar records.

Figure 11. Schema of a gemel window and a series of images of windows which respect that schema.

This is implicit in the 3D model; algorithms should be used or developed in the inverse direction in order to recognize it (e.g. using machine learning techniques).

Some projects have conducted similar research; for example, in the project 'Indianamas'[19], a system was developed in order to recognize the subject of prehistoric rock carvings from their 2D schematic representation using an ontology as the base for their labelling (Mascardi et al., 2014).

[19] indianamas.disi.unige.it

2 THE LIFE OF ARCHITECTURAL HERITAGE INFORMATION

2.1 The production

In the framework of this thesis, the production and collection of thematic data about cultural heritage is not considered in detail, since it is, obviously, a very broad topic. Therefore, we refer to multidisciplinary documentation produced by competent entities, each in its own field.

The geometric documentation possibilities are instead here briefly described, since they are to be included in the information about architectural heritage and it is useful to know some of their characteristics. The acquisition methods permit the obtaining of a large amount of data in little time, with high levels of detail and resolution. This information should not be lost in the representation; therefore, the data models and the reference ontologies should be adequate.

The characteristics of cultural heritage impose particular requirements on the surveying methods.

Large representation scales (up to 1:10 – 1:20) can be required. This implies that the achievable accuracy, reliability, and data density could be decisive to effectively survey the needed information. Some guidelines about these requirements are given in the document (Bryan, Blake, 2000).

Other appreciable properties of the techniques used for the generation of 3D models are (Remondino, Rizzi, 2010):

- *portability*: since many sites have limited accessibility or unsuitable conditions, it is preferable if the instruments are easy to be transported and can function autonomously;
- *possibility to acquire the data remotely*, so that the studied object is not touched avoiding possible damage;
- *fast acquisition and processing*: it is preferred to require limited time for acquisition for both not interrupting further works or the normal life of the studied object, and permitting a quick acquisition in case of particular vulnerability or, even, emergency;
- *flexibility*: due to the great variety and size of sites and objects, the technique should allow different scales and it should be applicable in any possible condition;
- *low cost*: the often limited budgets make low cost a necessity, besides always being a preferable condition for economic sustainability.

For responding to these requirements and to adapt to the different conditions imposed by the individuality of each surveyed object, it is necessary to be aware of the potentialities and characteristics of the available instruments and techniques. Moreover, an integration of the data deriving from the different methods is often required in order to exploit the possibilities of each technique and finally to produce a complete and improved model. Some examples of similar practices can be found in (Bastonero et al. 2014; Lerma et al., 2010; El-Hakim et al. 2004; and additional references reported in bibliography).

Common digital measuring sensors can be classified as ‘range sensors’ (Remondino, 2010), also called ‘active sensors’, since the impulse is given by the instrument and the response of the object is measured, and ‘passive sensors’, or ‘image sensors’, which record the light (in its spectral components) coming from the object. In the following, some of the most common techniques are described using active methods (the laser scanning technique) and passive devices (using images as the base, photogrammetry integrated by structure from motion contribution is presented).

All the techniques have somehow changed the critical workflow of the metric recording: they permit the efficient management of large quantities of data in little time. Therefore, the traditional surveying workflow followed, for example, with the topographic method (object interpretation – interesting limited numbers of feature individuation – feature measurement and plotting), has changed to the rapid measure of a high number of superabundant data, which are in a second step interpreted and chosen. This is due to the low cost and the velocity of new acquisition methods and measurement devices.

2.1.1 The survey techniques

A number of methods exist to survey architecture. However, only some of them are reported in the next sub-sections, which are chosen for being the more investigated ones in the course of the studies preparing this thesis. Moreover, they are the ones used for the survey of the case study employed for testing the proposed system.

2.1.1.1 LiDAR techniques

LiDAR techniques (Light Detection and Ranging) are well affirmed in cultural heritage application fields. The technology was developed during the last half of the 20th century (Abdel, 2011) and in the first decade of the 2000s developed to become a technology within everyone's reach. The accuracy and reliability improved, the prices became more affordable, and the reduced size and increased velocity in measurement made the Terrestrial Laser Scanner (TLS) suitable to record cultural heritage documentation also in more difficult circumstances (Bonfanti et al., 2013; Fröhlich, Mettenleiter, 2004; Barton, 2009; and others in the list of references). It is also the most affirmed technique for the survey of interiors.

The laser produces an optical impulse towards the object and measures the difference between some parameters in the emitting phase and in the return phase, thus establishing the distance of the reached object.

Different kinds of laser scanner exist; in particular, they can be divided by the kind of device in:

- *airborne* laser scanners, used mainly for cartographic aims;
- *terrestrial* laser scanners, largely employed in built environment;
- *hand-held* laser scanners, which are useful for closer measurements, which permit to reach a major detail.

They can be classified by the measuring principle in:

- *Time-of-Flight Laser (*TOF) – The distance is established by measuring the impulse time of flight to the object and return;
- *phase-shift laser* – it measures the shift between the outputted wave and the one that is received back;
- *triangulation-based systems* – they use the same measuring principle of photogrammetry: the point is measured as the intersection of two straight lines with known direction in the space.

The quality of the measurement, besides the nature and quality of the sensor, depends on the nature of the reached surface. In particular, the reflectance and luminosity properties can affect the measurement, and some reflective surfaces such as glass or water can cause problems in the measurements.

2.1.1.2 Imaging methods: photogrammetry integration with computer vision techniques

Photogrammetry is not a new technique in acquiring cultural heritage measures. However, in the last decade some algorithms developed by a different discipline, computer vision, have been integrated in the processing by several forms of software, and the great potentiality of the method allowed the image-based processing to be largely employed again (after a period in which the laser scanner seemed to be almost preferred for acquiring more points in less time).

The discipline of computer vision contains several research branches (Szeliski, 2010). One of these is structure from motion (SfM), which can reconstruct the sparse 3D point cloud of an object starting from a set of overlapped images (Snavely, 2006). The measuring principles of photogrammetry are implemented using different algorithms to generate a 3D model of the object in the processed images. Then, the photogrammetric technique of bundle block adjustment is used to improve the measurement quality of the orienting and calibration results. The process is automated with algorithms for the automatic tie-feature extraction and dense cloud generation. The resulting tools offer powerful techniques for obtaining dense, reliable and accurate 3D models in a short time using cheap measurement devices.

An extensive literature also exists about these methodologies (some studies are reported in the reference section).

From the cultural heritage point of view, the method offers great advantage for the achievable high density of points and the high quality of textures and orthophotos that can be produced when using high-resolution images as input files.

Moreover, the method is sufficiently flexible to also permit the use of images acquired for different aims, possibly in a time in which some information (that can now be lost) was visible (Snavely et al., 2006). This can be important for the *a posteriori* documentation of some no longer visible stratigraphy or some objects photographed in the past and then destroyed (Gruen et al., 2004).

The method also permits the use of images from low-quality cameras, such as action cams or similar. This can be useful for using data acquired at minimally accessible sites or for different aims than 3D reconstruction. In summary, this possibility increases the quantity of data useful to extract information.

2.1.1.3 UAV perspective

The recent spread of the use of unmanned aerial vehicles (UAVs) is given by the privileged point of view they offer for acquisitions. The possibility to reach a nadir close range point of view on the sites (especially in comparison with the traditional airborne acquisitions from planes or helicopters), the remote acquisition and control of the flight (without touching the object), and the low cost, quickness and high resolution of the acquisition make the acquisitions from UAVs a strategic method for surveying the built cultural heritage.

Many kinds of sensors can be mounted on the UAVs, depending on the survey aims: high-resolution cameras, LiDAR sensors, multispectral cameras, thermal sensors, video cameras, and any kind of sensor compatible with the dimension and payload requirements can be useful to use.

A further technique in the use of UAVs for built environment documentation consists in the oblique acquisition of images, so that the façades of buildings can also be recorded from the sky. Besides the velocity considerations in acquirement, the remoteness of the operation can also represent a great advantage when terrestrial accessibility is limited.

A broad literature also exists about these methodologies (some studies are reported in the list of references).

2.1.2 Georeferencing

The cited techniques permit the production of 3D models with very high detail and accuracy in the surveyed geometry. However, some measured reference is necessary for giving the models a metric value (scale and orientation) and locating them in a common reference system, which permits the consideration of the represented object in its context and the analysis of its spatial relations with other objects (e.g. Brumana, Achille, 2007). This permits the advantage of having a holistic view of the studied site.

Particular attention was paid to the problem of georeferencing cultural heritage when dealing with comprehensive GIS (e.g. the Italian case of the Risk Map). In these circumstances, some guidelines were investigated about georeferencing methods for cultural and architectonical heritage (Brumana et al., 2005; Brumana, Achille, 2003). Moreover, some criteria for archiving georeferenced data in the catalogues and defining the suitable kind of georeferencing are defined in Italian inventories (Section 6.2.6) guidelines (Ventura, 2014). They deal with the possible kind of located place regarding a cultural heritage item (classified as: storage; manufacturing; exhibition; origin; or retrieval place). Other specifications regard the georeferencing as an approximated point or exact position and the technique used to georeferenced the data, including the possible survey methods, some of which were described in Section 2.1.1.

Another issue often treated together with the georeferencing problem regards the kind of geometry being georeferenced: a point, a line or an area, plus the 3D models following more recent systems capabilities. This can influence the georeferencing results, especially for particularly small (e.g. museum objects) or large objects (whose location can considerably change if represented as a point or as an area).

A further topic essential for architectural heritage documentation is the representation of vertical surfaces, such as façades. This can be difficult in a unique 2D system, like the cartographic one. Some representation environments, more amenable to 3D representation, can change the reference system for additionally adequately managing the information related to such planes in their own local coordinate system; e.g., Computer-Aided Drawing (CAD) systems or Building Information Model (BIM) can manage this. Instead, in GIS only cartographic reference systems are admitted. A variety of solutions are proposed in some of the examples presented in Chapter 3. However, the possibility to manage both 3D and 2D representations and to pass from one to the other without losing the reciprocal reference could be a challenging issue.

Furthermore, a large fraction of cartographic disciplines is dedicated to the definition of reference and projection systems, the representation deformations they involve, and the methods of transformation from one system to the other. In the past, for architecture 3D modelling, often a local reference system was sufficient, but this precludes the possibility to exploit the advantages of an integrated management and harmonisation of different kinds of information and the consequent powerful tools. The possibility to share the produced information on the internet cause the georeferencing to a known and common reference system[20] to be a necessity, since it is part of the disambiguation metadata for a correct data interpretation.

2.1.2.1 The Georeferencing techniques

As well as for the survey techniques, various instruments and methods are available for georeferencing the data. They can be chosen based on different criteria and the specific requirements of the surveyed site.

As explained in the previous section and reported in some reference guidelines for cultural heritage (e.g. Ventura, 2014), different kinds of georeferencing can be employed, carrying different exigencies in the measurements and influencing the choice of the technique to be used.

The requirements to consider for selecting the most suitable technique are:

1) the required precision;

existing instruments can reach different precisions, from a few millimetres to a few metres. It is therefore important to know what is the target accuracy of the survey and the georeferencing, also considering the characteristics listed in (Ventura, 2014); that is, if the aim is the plotting of a point as the approximated position of a monument as a location in a regional-scale map, very sophisticated instruments are not necessary;

2) the rapidity of the measuring phase;

some techniques permit faster measurements, but some conditions have to be verified to avoid negatively influencing the quality of the measurement;

3) the costs;

different instruments (and connected processing software) have different costs that should be evaluated for each case, depending on the available budget, the required quality, and further reasons;

4) accessibility;

the nature of the site and the position and accessibility of the points to be measured are also to be considered;

5) presence of needed infrastructure;

a determining factor for using some technologies is the availability of some supporting infrastructures, which is to be evaluated for each specific site.

An evaluation of these factors can yield the selection of a specific technique or, more frequently, of some techniques, integration to reach the best result possible for the specific needs.

The two more used techniques for georeferencing the models are GNSS instruments and technologies (Section 2.1.2.1.1) and the Total Station (Section 2.1.2.1.2), often with an integrated approach. They

[20] http://spatialreference.org/
https://en.wikipedia.org/wiki/Spatial_reference_system
http://www.crs-geo.eu

are employed in the acquisition and processing phases for measuring the coordinates of some reference points to be used for the processing and orientation of the 3D models.

The measured points consist generally in Ground Control Points (GCP) that can be purposely marked with expressed devices or identified on the object as ‘natural points’ (existing features well-defined on the object that can be used as reference). Their coordinates are computed through the different techniques and are used as references to be collimated during the processing phases of the 3D models, or even in some *a posteriori* georeferencing activity (Bitelli, Gatta, 2012). For better refining the model and optimizing the georeferencing, the GCPs should be superabundant in number (more than three references, which are the minimum, are needed, and generally many more are required and preferred for correctly and effectively estimating the residuals on each point). It is also important that they are well-distributed on the object, for this reason it is critical to consider the possibility of integrated techniques for their measurements.

Other methods for georeferencing the models through known measured coordinates in an absolute reference system include the direct photogrammetry, which can be extremely useful in emergencies and for aerial acquisitions (Chiabrando et al., 2013). In this case the measured coordinates refer to the position and asset of the photographic shoot for each image, and can be used as reference for aligning the images when building the model and orienting it directly in the chosen reference system.

2.1.2.1.1 GNSS positioning

Different georeferencing techniques exist, but, generally, all of them refer to positions measured with GNSS methodology (Global Navigation Satellite System) (Cina, 2014).

GNSS positioning works with a spatial distance intersection of the measured signals emitted by some satellites orbiting around the Earth and belonging to different constellations:

- the American NAVSTAR GPS (NAVigation Satellite Timing And Ranging Global Positioning System);
- the Russian GLONASS (GLObal NAvigation Satellite System)
- the European GALILEO;
- the Chinese BeiDou or COMPASS;
- the Japanese QZSS (Quasi-Zenit Satellite System).

The measured data permit the use of a global reference system, which can be considered when publishing the data on the web. Each constellation has its own reference system, but the relations among them are known and an easy translation is possible. Moreover, given that the signals of the satellites are similar, they can be used contemporarily in a unique GNSS. The measures referring to different reference systems (such as local or national reference systems and so on) can be translated into another one using known relations (e.g. Manzino, 2001). The most used reference system for publishing data on the web is WGS84 (World Geodetic System, 1984), which is the reference system of the GPS constellation (Cina, 2014). It is, at present, the most implemented one in the web platforms (for example, Google Earth uses it).

The very widespread code for defining reference systems univocally in applications (both web-based and desktop), as well as in GML files (Section 6.3.3) is the European Petroleum Survey Group code (EPSG). It identifies the International Association of Oil & Gas Producers’ (IOGP’s) EPSG Geodetic

Parameter Dataset: a collection of definitions of coordinate reference systems and coordinate transformations, which may be global, national, regional, or local in application[21].

The points measurements by means of GNSS methods can reach different qualities and accuracies depending on the technique and instrument employed. A series of devices (having different costs) can measure GNSS position, obviously with different accuracies: from palmar and low-cost receivers, with prices around 100€ (such as the ones integrated in smartphones), which can measure with metre error to very sophisticated antennas achieving millimetre accuracy.

Furthermore, other issues have to be taken in account: for example, the measuring technique (Cina, 2014).

The GNSS positioning can be:

- absolute (or stand-alone), where the coordinates of a point are determined in a 'global' reference system;
- relative, in which the components of the vector linking two points (baseline) are computed; in this way, the systematic errors (bias) affecting the range of the two stations on each satellite are reduced or eliminated. Some permanent stations on the Earth's surface continuously receive the signals and, through techniques that are objects of a specific branch of geomatics, permit the definition of their position in time and the optimisation of the measure of the antennas positioned on the reference points of the topographic network for a detailed survey;
- differential positioning, similar to the absolute one, but executed by correcting the range satellite – receiver with a differential correction computed by a 'base' station. As in the relative positioning, systematic errors are reduced or eliminated.

Especially for establishing which mode among these can be considered, it is necessary to know if a permanent-station infrastructure supporting the measures is available at the site.

Moreover, the measures are divided into:

- static, when the receiver is positioned on each point for some time (at least a duration of several minutes); this is used, for example, when measuring the network vertices;
- kinematic, in the case that the receiver is continuously moving (a few seconds are needed for recording the measure; this is the case of the rapid measurement of GCPs on the ground in wide areas, by means of the Real Time Kinematics (RTK) technique).

The positioning can moreover be realized:

- in the post-processing phase (the data are processed after the acquisition);
- in real time (the position is immediately available during the measurement campaign).

The evaluation of the needs for each case and the check about the availability of supporting infrastructures should guide the choice of the method.

Some critical issues can be identified related to using this technique. The most important one regards the possibility that the satellite signal cannot reach the site, for the lack of communication (generally, the GNSS positioning uses radio waves, or the Global System for Mobile (GSM) Communication, networks). Another one is due to the structure of the site: the presence of obstacles can intercept the waves and disable the positioning (e.g. trees, indoor environments, urban canyons, etc.) (Figure 12).

---

[21] www.epsg.org

An advantage of the GNSS measurement can be the fact that the measure of each point is independent of the others and the possible errors are not propagated.

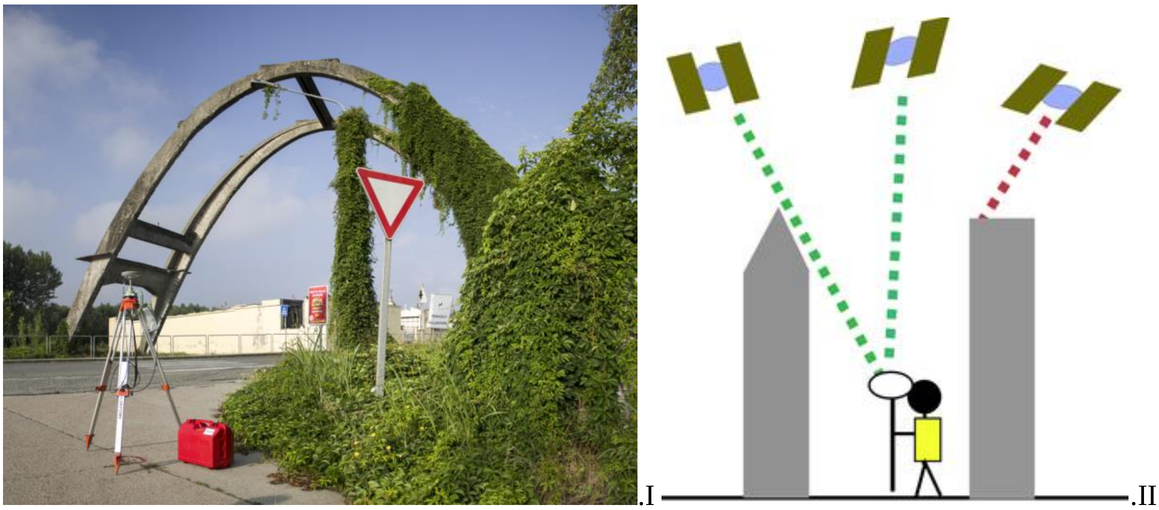


Figure 12. The GNSS receiver can suitably measure with good visibility (I)[22], while their performance are lower in presence of some obstacles (II)[23].

The GNSS positioning refers to an abstract ellipsoid. This has the consequence that the height measurements must be corrected when considering the geoid reference (Figure 13).

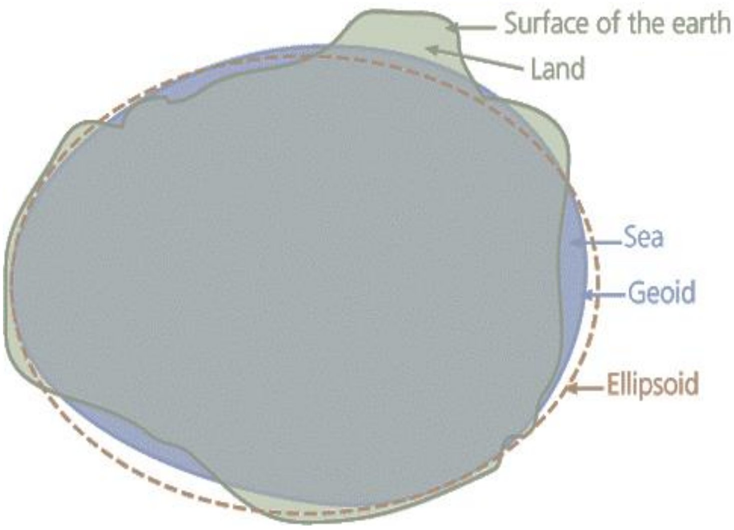


Figure 13. Geoid vs ellipsoid[24]

[22] GNSS receiver acquiring during the measurements realized by the student Team DIRECT (funded by the Politecnico di Torino) in a stage in Morano sul Po, Italy (http://www.5xmille.polito.it/progetti/2014/direct_pronto_intervento_per_la_documentazione_in_caso_di_catastrofi)

[23] http://www.mdpi.com/2220-9964/4/4/2769/htm

[24] http://www.esri.com/news/arcuser/0703/geoid1of3.html

2.1.2.1.2 The Total Station

For integrating or thickening the measurements, the Total Station (TS) is used: it permits the collection of points that are difficult to be measured with GNSS techniques for accessibility reasons (such as the points on a façade), or due to the absence of GNSS signal (absence of the communication networks and structures or presence of obstacles). Among these cases, a specific problem regards the indoor environments, for which the integration of the techniques is essential in order to achieve a georeferencing of the object in a global system.

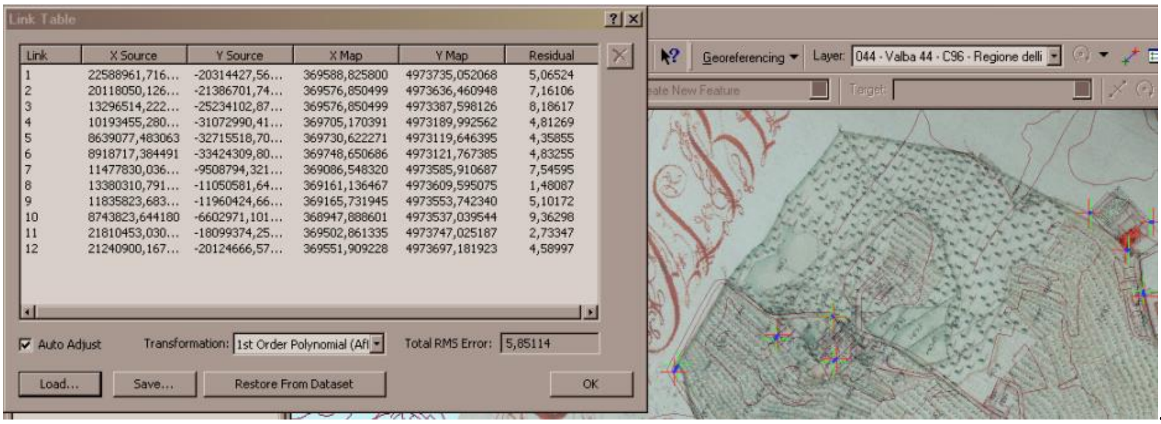
The advances in the instrumental technologies permit the existence of complex total stations, which improve the integration among different techniques. More developed Total Stations, in fact, permit users to:

- measure traditional angles and distances for computing points coordinates having a reference network or origin point, as was the original and remains the main aim of the total station;
- measure its position through the integration of GNSS receivers;
- use integrated GNSS receivers for linking to a GNSS network and possibly enhancing the surveying at further sites (e.g. indoor or covered areas);
- exploit some functions like the automatic acquisition of measures, which is similar to laser scanning, although having lower performance (especially in terms of acquisition time); however, it could be useful in some cases;
- recording images: this permits the user to either employ the images for photogrammetric aims or for storing them as reference as a sort of eidotype about the measured points (this is useful especially for documenting natural points measurements, very useful in cultural heritage contexts where accessibility is reduced and the objects can be very complex).

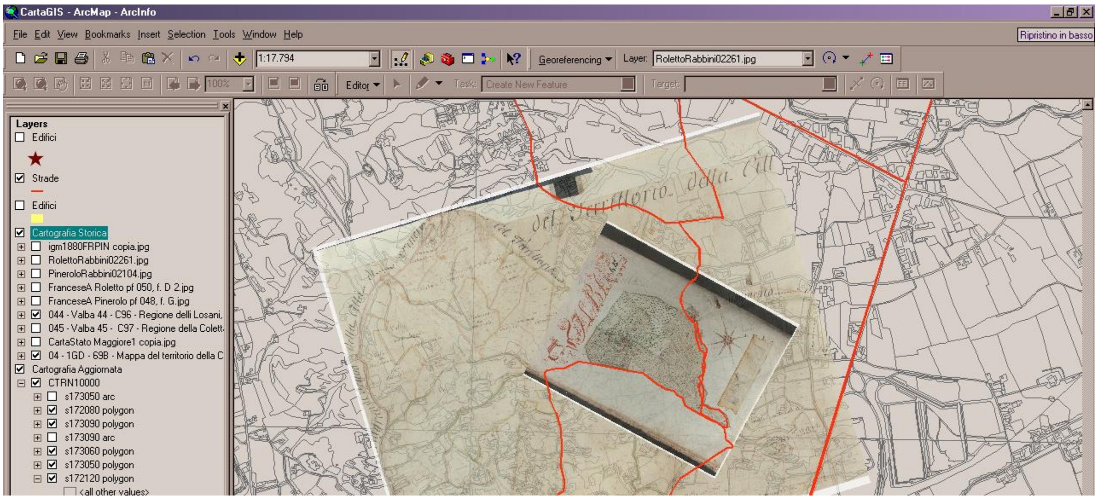
In summary, considering the integration among techniques is always recommended.

2.1.2.1.3 Further georeferencing methods

Other techniques can be used for georeferencing aims. A particular case involves the georeferencing procedures to use if the position has not been measured on-site. One of them consists in the application of homographic transformations to superimpose a raster product to some georeferenced references (e.g. Brumana et al., 2005). This is a spread method for the geolocalisation of historical digitized maps (Figure 14).



.I



.II

Figure 14. Screen-shots of the georeferencing process of some digitized historical maps by means of a homographic transformation computed basing on the collimation of some reference points (I) on a georeferenced and updated map in GIS tools (ESRI ArcGIS) (II) (Noardo, 2012).

Other more innovative methods can be considered for the automatic or semi-automatic registration of raster images acquired with different spectral band values, using image analysis applications and computing environments (Aicardi et al., 2016).

2.1.3 Metric products for Architectural Heritage

From the data acquired with the cited methodologies, different kinds of products can be extracted or processed depending on the specific needs. The more used kind of data are summarised in the following. They are the data that can be structured in systems for the representation and sharing of heritage information.

2.1.3.1 Point clouds

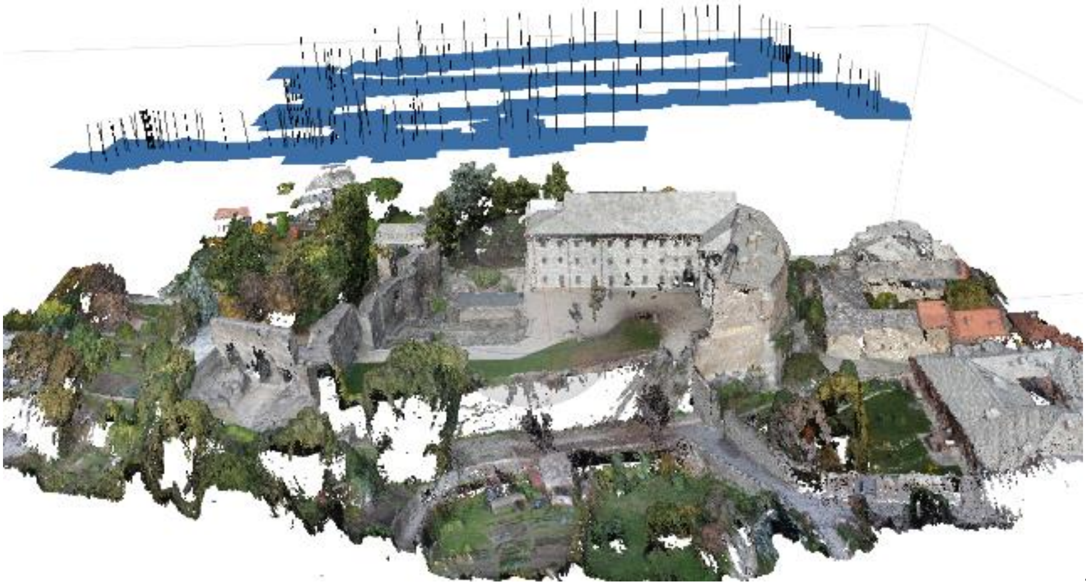
A point cloud is one of the first processing products of a survey with the examined techniques (Figure 15). It is the set of X, Y, Z coordinates in a reference system that describe the position in the space of a large number of points belonging to the object. The previously described techniques permit very dense point clouds (as well as millimetre scale mutual distances). The 3D model that the whole cloud constitutes has the advantage of not being affected by the approximations present for example in interpolation products. They can often be further optimised, filtered and processed in order to reduce the noise, which can affect them, and improve their quality.

They can be used as base for further processing steps:

- interpolation of raster elevation models and 3D triangulated models;
- profiles extraction;
- base for interpretation, classification and analysis for specific issues.

Some rules about the quality, density, and accuracy of point clouds can be found in specific documents (Brovelli et al., 2009; Bryan, Blake, 2000).

.I



.II

Figure 15. Examples of point clouds produced by means of TLS (I), the hall of honour of the Stupinigi royal residence; and from structure from motion processing of images acquired from UAV (II), roman acropolis of Susa in Italy.

2.1.3.2 Orthophotos

An orthophoto is an image whose perspective deformations are rectified. It is therefore a metric product enriched with the possibility to read some thematic information from it, which could also be only human-interpretable, since it is still unclassified.

The support of orthophotos for cultural heritage is important to critically read surfaces and consequently map them, avoiding approximation errors (present, for example when mapping features on an orthographic drawing basing on distorted photos or on-site sources). The orthophotos can also be used to plot traditional 2D drawings (Figure 16), which are still the base for the restoration projects and which can enlighten some important building characteristics.

The availability of an uninterpreted image of the state of fact of a monument is important to be archived because it represents a basis for being read in the future, for monitoring reasons, or for analysing different aspects not considered before (Figure 17).

The orthophotos can be classified following many criteria (Brovelli et al., 2009; Bryan, Blake, 2000).

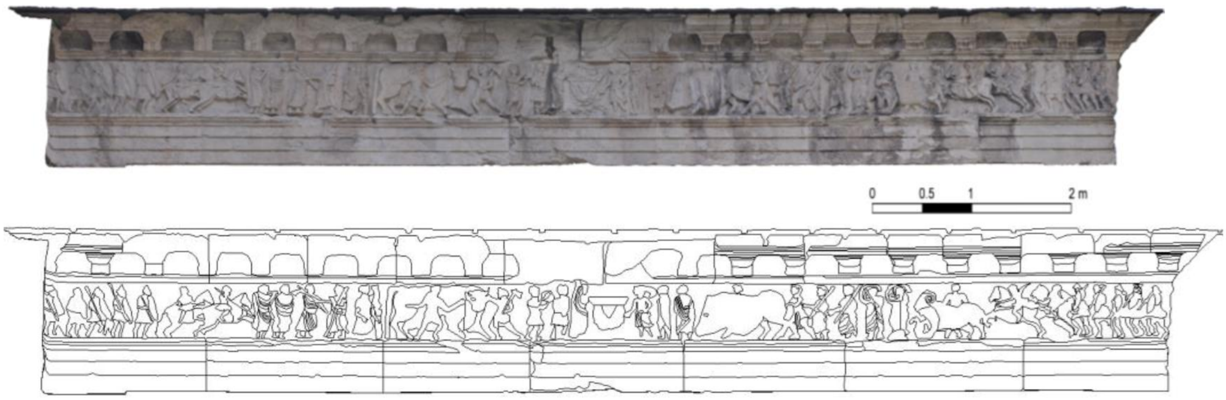


Figure 16. Orthophoto and 2D plotting of the Susa Arch of August frieze (Chiabrando et al., 2015).

Figure 17. Orthophoto of the intrados of the balcony of the Stupinigi royal residence hall of honour, generated by SfM methods.

2.1.3.3 Digital Elevation Models

Digital elevation models (DEM) are points describing the height values of a surface in a selected direction (for example, the nadir direction). Usually these are regularized in a regular grid for simplifying the result and are often translated in raster formats, in which a pixel corresponds to the point of the regularized grid. The height value is stored in the pixel.

Particular elevation models are the ones referring to the nadir point of view, used in cartography. These are classified as digital terrain models (DTMs), which represent the elevation of the terrain without objects on it, and digital surface models (DSMs) which include the representation of the objects above the ground (trees, buildings, bridges, etc.) (Figure 18). In architecture analysis, it can also be useful to consider an elevation model of a façade or a ceiling, for verifying the direction of the surface and possible deformations. In this case, it is called simply ‘DEM’ (Figure 19).

Elevation models are useful to analyse the surfaces also exploiting automatic algorithms integrated in common GIS or raster analysis software, which are generally faster than similar ones operating on vector formats. This permits the visualization of the eventual irregularity of a surface or can serve as the base for GIS spatial analysis.

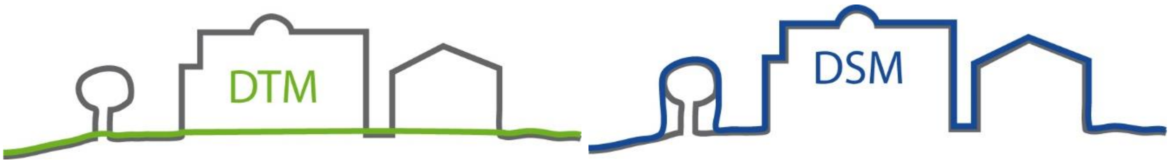


Figure 18. Schemas of data included in DTM and in the DSM.[25]

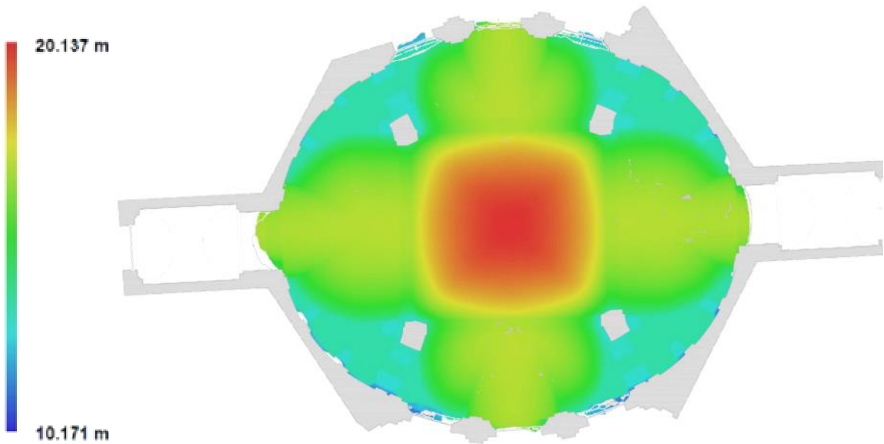


Figure 19. Example of DEM (Stupinigi Hall of Honour vaults)

2.1.3.4 Traditional drawings in orthogonal projection

Even if the impact of 3D visualisation is high, and the analysis of a 3D object can offer improved potentialities, 2D drawings still remain the basis for understanding certain building features. Some building techniques, proportions, and generating geometry effectively emerge from an orthogonal 2D drawing. Moreover, these drawings are often still the basis for the redaction of intervention projects.

[25] http://eijournal.com/print/articles/global-elevation-data-enhance-exploration-and-development

The new survey techniques can help in producing similar outputs by offering a plotting basis or automatic profile extractions (Figure 20 - Figure 21).

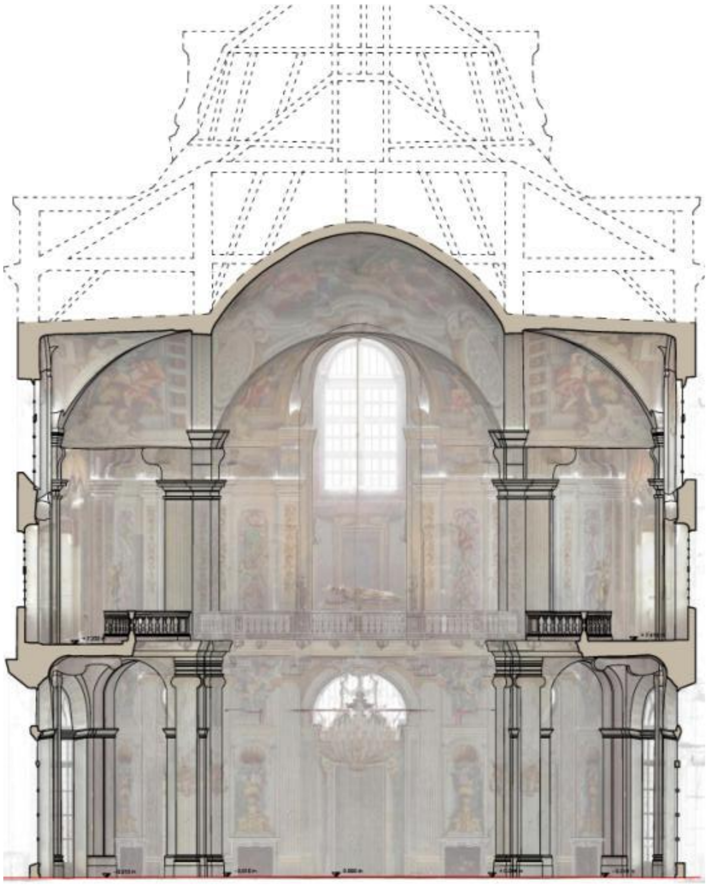


Figure 20. The section is plotted using a TLS survey results as base (projected as a texture in the back). The object is the Hall of Honour of the Stupinigi royal residence.

Figure 21. The orthophoto reproducing the vaults was generated by digital photogrammetry – sfm image processing.

2.1.3.5 3D representation objects

(Hichry et al., 2013) classifies the 3D representations, which are generally derived from the previously described products, as follows:

1) Parametric vs non-parametric representation.

Parametric representation describes the model using a set of parameters such as the height, the length, the radius, etc. (A cylinder is described along its axis and its radius). Non-parametric representation uses other ways of characterization such as triangular meshes.

2) Global vs local representation

In a global representation, the entire object is described while in a local one only a portion of the object is characterized.

3) Explicit vs implicit representation

The explicit representation allows a direct encoding for the shape of the object (e.g. triangular meshes), whereas the implicit representation allows an indirect encoding for the shape using an intermediate representation (e.g. a histogram of normal surfaces).

Among the various possibilities, two options are mainly used for representing architectural heritage: meshes and B-Rep.

A mesh is a triangulated surface generated by the interpolation of the measured points. It is a 3D non-parametric explicit representation. The advantage is given by the small approximation to an ideal geometry of the representation: the measured points are simply interpolated following variable parameters and algorithms, but the surveyed object complexity and irregularity can be maintained.

The B-Rep describes shapes using a set of surface components that constitutes the surface limits. Volumetric representations describe shapes with geometric solids known as Constructive Solid Geometry (CSG), which consists of building complex shapes starting from simple geometric primitives (such as cube, cylinder, sphere…) by combining them using Boolean operators like union or intersection. Compared with the B-Rep, CSG are more intuitive but are not as flexible because of their limited library of primitives. This representation is mainly used in 'as-built BIM' (Hichry et al., 2013). Generally, with the as-built BIM approach implicit representations are preferred, which could be effectively segmented by automatic algorithms and have generally a local interest.

In this research, non-parametric, global, and explicit representations are used. These permit the storage and documentation of the details on the state of fact of the building avoiding any kind of approximation and permitting their contextualisation in a global frame. A texture can enrich 3D representations (Figure 22).

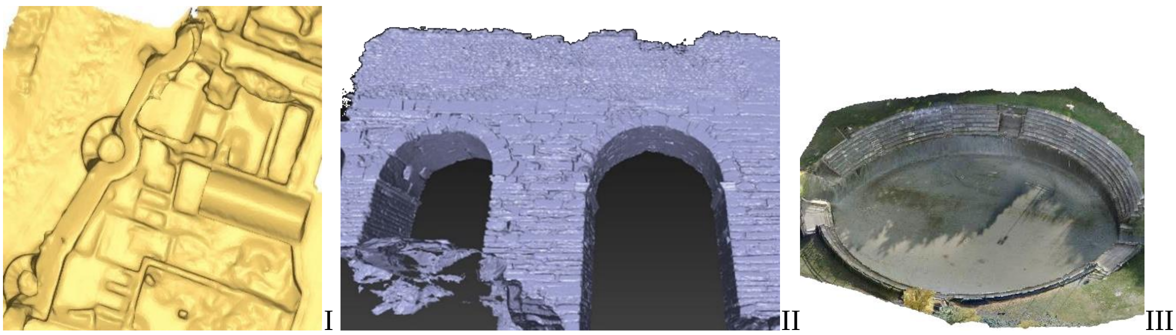


Figure 22. Examples of meshes: (I) Susa (Italy) acropolis (UAV - SfM survey); (II) Susa acqueduct arches (TLS survey), (III) Susa amphitheatre (UAV – SfM survey) with texture.

2.2 The management

Large volumes of information can be stored in the described collected data. However, if they remain disconnected from each other and they are not managed in order to effectively represent, analyse and communicate the information they carry, their meaning will be blurred or useless, as could happen if some important document is hidden in a drawer and its context is forgotten.

One of the systems to effectively manage the data is described in the following sections. It is the use of a Geographical Information System (GIS), which permits the archiving of the data in a unique frame for having advantages over more complex systems.

Moreover, a process that can be extremely useful if applied to the very large amount of data about cultural heritage is data mining, which is the process of discovering valuable information and meaningful patterns within large data sets (including spatial data). The difference from a simple database query is that users may not know what information or patterns they are seeking in advance (Worboys, Duckham, 2004). This is possible if the data are properly archived. Furthermore, recently developed technologies can extract information from very large quantities of data, and exploit different repositories (such as the web). The eventuality to apply this process is to be considered for the goals of this thesis.

2.2.1 A representation for integrated management: GIS

An information system is an association of people, machines, data, and procedures working together to collect, manage, and distribute information of importance.

A geographic information system is a computer-based information system that enables the capture, modelling, storage, retrieval, sharing, manipulation, analysis, and presentation of geographically referenced data (Worboys, Duckham, 2004).

The great advantage of GISs is the possibility to manage heterogeneous data in a unique frame: multi-source and multi-formats data (both raster and vector formats for geometry data and a variety of types, including raster image type or further multimedia products, for thematic alphanumerical data). Moreover, dynamic data can be stored and analysed. Essential to the interpretation of the archived information are metadata, which can be stored in GIS associated to each data.

GIS processing potentialities permit the visualization of thematic georeferenced data, to select and retrieve specific information, to compute statistics about data, to compare data, and to generate further data by combining and analysing the already stored information.

The functions to be provided by a GIS according to (Laurini, Thompson, 1992) are the following:

- *provide tools for the digital representation of spatial phenomena (data acquisition and data encoding);*
- *handle and secure the encodings efficiently by providing tools for editing, updating, managing and storing and for converting, verifying and validating those data;*
- *foster the easy development of additional insight into theoretical or applied problems, by providing tools for information browsing, querying summarizing;*
- *provide facilities for analysis, simulation and synthesis;*

- *assist the task of spatial reasoning by providing efficient retrieval of data for complex queries;*
- *create people compatible output in varied forms*.[26]

Two main aspects characterize the GISs. On one hand, they manage spatial analysis (Fotheringham, Rogerson, 2013; Maguire et al., 2005; Goodchild, 2004), which exploit the georeferencing of the data for extracting further information (buffer, overlay functions, vicinity analysis etc.). A large part of this field regards the surface analysis, which deals, for example, with elevation rasters in order to perform hydrographic analysis (basins, watershed, etc.) or solar radiation analysis (basing on aspect and slope characterisation of the surfaces), or visibility analysis. A number of algorithms are integrated in GIS software for realizing these issues.

On the other hand, huge potential comes from the structuring of the archived data in databases. This permits the enrichment of the previously produced data with some context (metadata about their production and their memorization, relationship with other data, conceptual model structuring them, etc.). In this way, a more effective form of information is produced (Worboys, Duckham, 2004). Moreover, the material about the criteria and procedure used for the definitions and demarcations of the data (metainformation) is stored.

Nevertheless, GIS enables all the procedures for data retrieval and computing in databases. For example, the execution of queries and selections, statistics about data, the generation of reports or graphs, thematic visualisations, etc.

2.2.2 Database

A database (DB) is a repository of data that is logically related, but possibly physically distributed over several sites, and required to be accessed by many applications and users. A database is created and maintained using a general purpose piece of software called a "database management system" (DBMS). For a DB to be useful it must be:

- *reliable (able to offer a continual uninterrupted service);*
- *correct and consistent (data items should be correct and consistent with each other);*
- *technology proof (it should evolve with each new technology development);*
- *secure (it must allow different levels of authorized access).*[27]

The definition given by (Worboys, Duckham, 2004) addresses with some implementation characteristics of a good database, which are the objects of study of different disciplines, such as Information and Communication Technology (ICT). A good database project, which is the preliminary phase for a GIS building, must keep them in mind, in order to develop application-independent schemas, and to provide explicitly all the information that can be useful in the future implementations or data interpretation.

Databases offer a number of tools for retrieving information from the archived data and presenting it for communication purposes. The same abilities are transposed to geodatabases (geographical DB) (GDB), introducing the spatial aspect for added potentialities.

[26] The italic text cites (Laurini, Thompson, 1992)

[27] The italic text cites (Worboys, Duckham, 2004).

2.2.2.1 The database modelling process

The preliminary phase is the data modelling, that is, the choice of appropriate data model to manage the data about a specific application domain. This is a schematisation and formalisation process that starts from the analysis of the perceived reality of the application domain by humans to reach a machine-readable format, which implements it in computer applications. It is important that this is neither difficult to use, to avoid neglecting the human aspects of the information system, nor inefficient, to avoid neglecting the computational aspects (Worboys, Duckham, 2004). A model is an artificial construction in which parts of one domain (the source domain) are represented in another domain (target domain). The purpose of the model is to simplify and abstract away from the source domain.

A meaningful concept related to modelling is ‘morphism’, a function from one domain to another that preserves some of the structure in the translation. This can help in determining how closely a model can simulate the source domain and how easy it is to move between the two domains.

The successive modelling phases have been defined by the ANSI/X3/SPARC standard since 1975 (Laurini, Thompson, 1992) and consist in:

- *external model* (application domain model description): it uses the natural language (high-level language) and is an analysis of the interested part of reality to be managed in the GIS;
- *conceptual model*: implementation independent, it formalizes the previous model by identifying the concepts (entities or class of entities) and the relationships among them;
- *logical model:* it is tailored to a particular type of implementation. It focuses on how the system will implement the conceptual model (system design);
- *internal model* (also called physical model): it describes the actual software and hardware application in a low-level language (system implementation). The final implementation will contain the self-description of the system (encoding both data and structure of the data), from which it would be possible to deduce the modelling process phases.

The progression of the phases are represented in two different but equally meaningful schemas: the (Laurini, Thompson, 1996) effectively illustrates the modelling process, emphasizing the research of a consensus in the early abstraction phases (Figure 23); in the meantime, the so-called ‘system life-cycle’ in (Worboys, Duckham, 2004) emphasizes the evolutionary and iterative nature of the development (Figure 24).

Knowledge is a process of piling up facts; wisdom lies in their simplification.

Martin Fischer

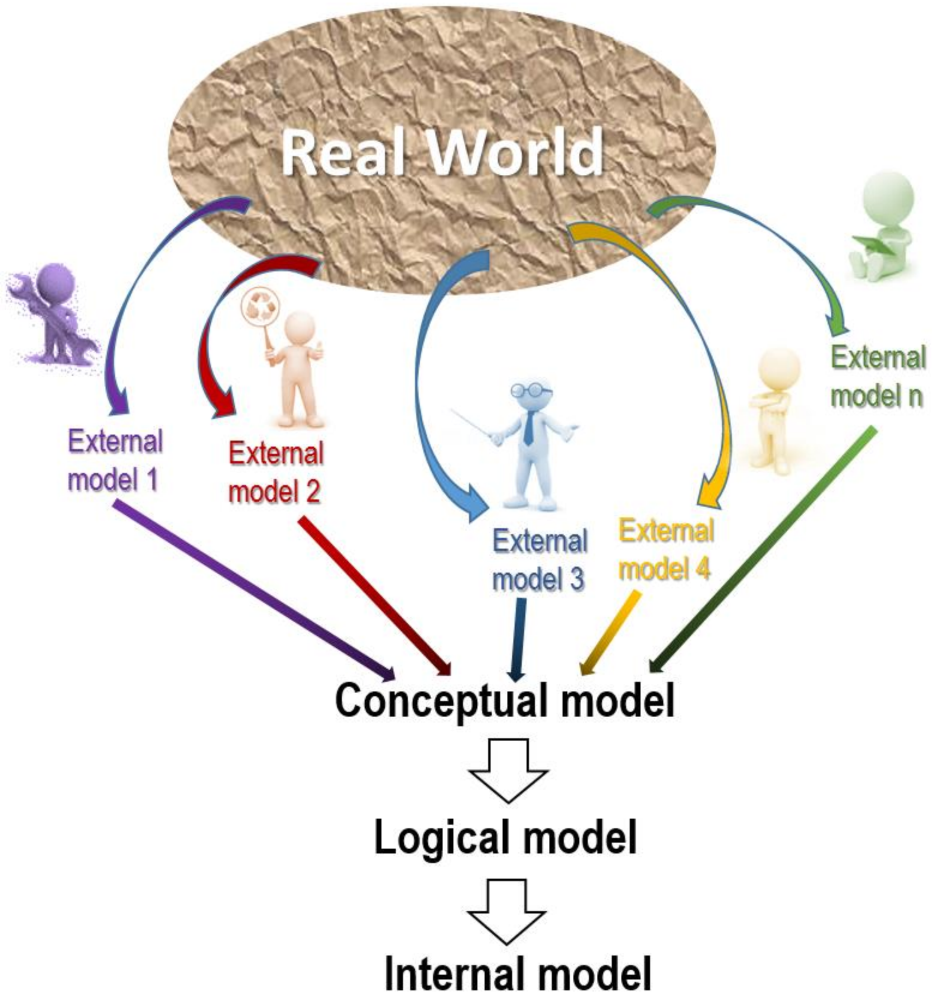


Figure 23. A framework for the design of an information system adapted from (Laurini, Thompson, 1992).

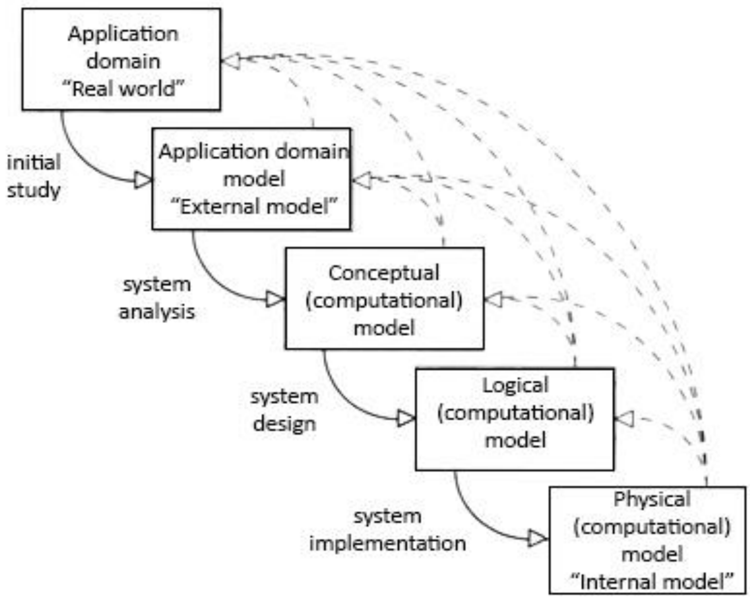


Figure 24. Modelling processes and the stages of system development underlying their iterative nature. From (Worboys, Duckham, 2004)

2.2.2.2 The conceptual model

Conceptual modelling is the activity of formally describing some aspects of the physical and social world around us for the purposes of understanding and communication.

John Mylopoulos

In the conceptual model, the entities to be managed in the database, their attributes and their associations are defined in a formal language. An entity is a phenomenon that cannot be subdivided into similar units (Laurini, Thompson, 1992). The conceptual model focuses on what the system will do (system analysis). It is the main disambiguation tool for the system and aims at an efficient storage and querying of the data. To this static component, often a dynamic component is added, which is related to its behaviour in operations (Worboys, Duckham, 2004).

Usually for modelling the database some standard notations are used, such as the entity-relationship (E-R) model[28] or extensions of it.

Among these, the Unified Modelling Language (UML) has some relevance (Section 2.2.2.2.1). It is used in the definition of standard data models, and therefore it is also employed in the application part of this thesis.

2.2.2.2.1 Unified Modelling Language (UML)

UML is a modelling language based on the object-oriented paradigm. Object-orientation is a specific approach for building a logic model, described later in the text (Section 2.2.2.5). However, it is used for conceptual modelling independently from the following translation to a chosen logical model.

It was defined in 1996 by Grady Booch, Jim Rumbaugh and Ivan Jabson at the Object Managemetnt Group[29] (OMG). The aim was to develop a unique language that could become a standard, as it happened later. It is now considered an industrial standard and it has spread in the informatics programming community as well as for conceptual modelling activities. In 2004, a new version (UML 2.0) was published, which includes newly organised elements and new tools[30].

Many kinds of diagrams can be built using UML, for programming projects, business management, and modelling issues[31]. For conceptual modelling, the UML class diagram is typically used. A synthesis of UML notations used in modelling is represented in Figure 25.

The Unified Modeling Language- UML - is OMG's most-used specification, and the way the world models not only application structure, behavior, and architecture, but also business process and data structure.[32]

[28] https://en.wikipedia.org/wiki/Entity%E2%80%93relationship_model
[29] www.omg.org
[30] https://it.wikipedia.org/wiki/Unified_Modeling_Language
[31] http://www.omg.org/gettingstarted/what_is_uml.htm
[32] The italic text cites http://www.omg.org/gettingstarted/what_is_uml.htm

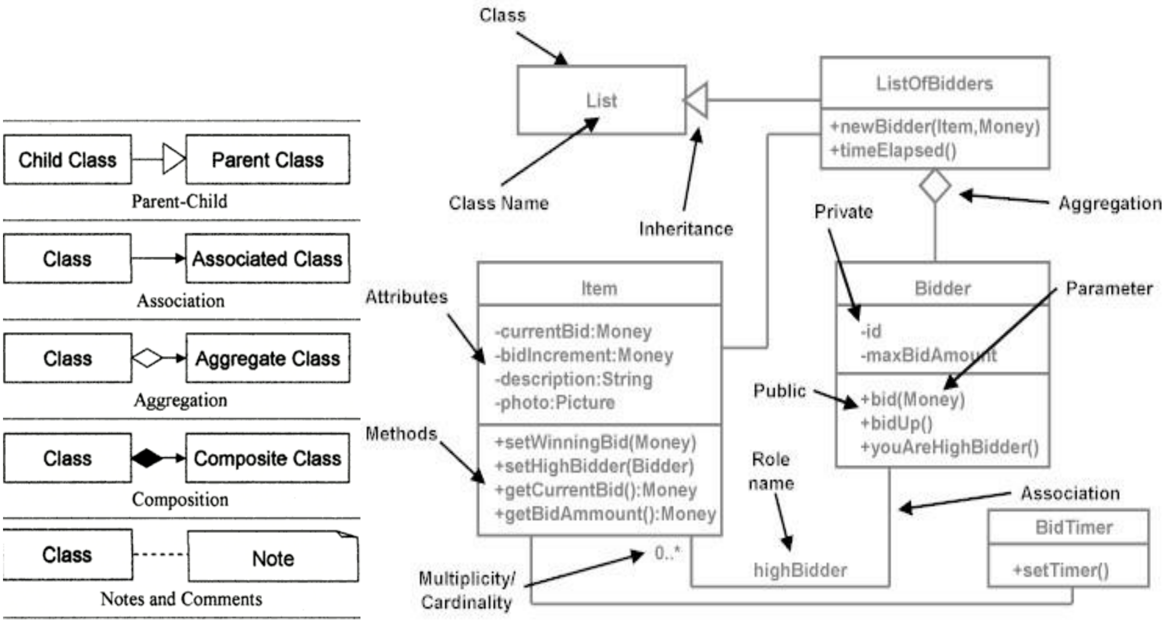


Figure 25. Examples of the most-used UML notations: types of association on the left, an example of class diagrams with labelled parts on the right. (Images retrieved on the web[33]).

An important consideration, which can be given in a UML object definition, is the stereotype: it appears above the class name like `<<stereotype>>`. It provides (also for translation procedures of the models) the interpretation of the class and consequently activates the specific translation or further operations. These are defined in an XML document called 'UML profile'.

UML supports stereotypes, which are an inbuilt mechanism for logically extending or altering the meaning, display, characteristics or syntax of a basic UML model elements.[34]

An extension of UML was proposed by (Guizzardi, 2011) for representing more specific mereological relationship among collectives and their subparts in conceptual modelling.

In UML, mereological association can be defined as 'aggregation', that is, the portion-object relation, or 'composition', that is used for representing the component – integral object relation (see the chapter below and Section 1.3.2).

2.2.2.2.2 Mereology for conceptual modelling

In conceptual modelling, mereological relations, detailed in Section 1.3.2, should be considered, especially for some aspects, for example transitivity (Guizzardi, 2009). Other rules that can be of interest for conceptual modelling can be stated, such as 'two components share a component only when one is a sub-component of the other'[35], and so on.

(Keet, 2006) reports the lists of part-of relations (listed below) defined by Odell, well known in object-oriented conceptual modelling. They are defined in Section 1.3.2, but are here synthetically reported:

[33] https://www.stickyminds.com/sites/default/files/article/2012/XDD6211imagelistfilename1.jpg
[34] http://www.sparxsystems.com/enterprise_architect_user_guide/10/standard_uml_models/stereotypedlg.html
[35] Slides Ghidini, C., 2012, Logics for Data and Knowledge Representation http://disi.unitn.it/~ldkr/ldkr2012/slides/OntologicalAnalysis1.pdf

- *component – integral object*;
- *material – object*;
- *portion – object*;
- *place – area*;
- *member – bunch*;
- *member – partnership*.

Some of them, such as the 'component – integral object' relation, can be represented, for example, in UML (Figure 26):

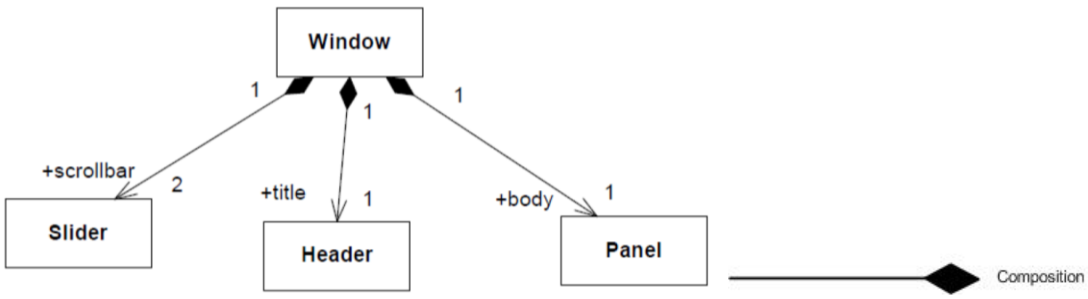


Figure 26. Composition representation in UML diagram

2.2.2.3 The logical model

In the logical model, the constraints and business rules of the systems are also explicated; among these, a system should be found for including the spatial business rules cited in the introductory part.

Many database architectures exist and have been defined, since the 1990s (Medeiros, Pires, 1994). Some of them are the hierarchical model; the network model, which, in addition to the hierarchical model structure, allows many-to-many relationships; and the deductive model, which allows the storage of rules for making inferences about the data. However, the most common ones are the relational model or the object-oriented model (Worboys, Duckham, 2004; Laurini, Thompson, 1992). A comparison of the two models and the analysis of the reasons for the two models and their respective success can be found in (Bloor, 2003).

2.2.2.4 The relational model

The relational model is built by the relationships between data items. A relational database is a collection of tabular relations (tables), in which the data are stored as a set of rows (the relations, also called 'tuples') in which some values fill in the attributes (labelling the columns). A relation is therefore a finite set of tuples (Figure 27). It is associated with a relation schema, which gives the structure of a relation and is a set of attribute names and the mapping from each attribute name to a domain (which establishes the format of the data contained in it) (Figure 28).

A set of relation schemes is the database schemes, while the set of relations is the relational database (Worboys, Duckham, 2004).

A primary key is chosen for uniquely identifying each record, and further integrity constraints can be set.

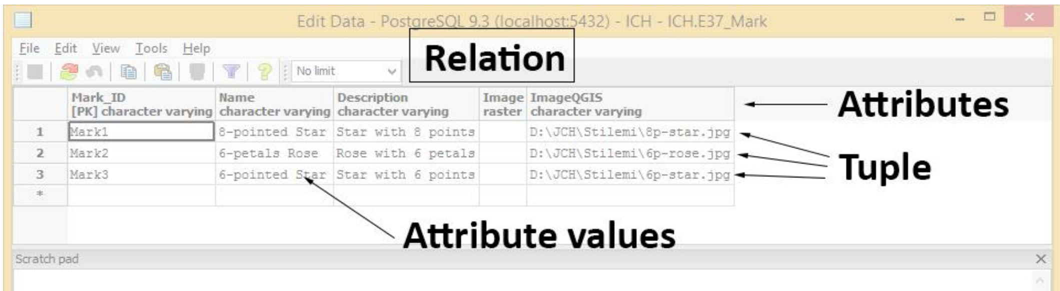


Figure 27. Example of a relation consisting in a table of a database.

Column name	Definition
<u>Mark_ID</u>	character varying NOT NULL
Name	character varying
Description	character varying
Image	raster
ImageQGIS	character varying

Figure 28. Example of a relation scheme. The underlined attribute is the primary key of the relation.

Moreover, operations can be performed on the relations. They are objects of the relational algebra. Some of the most common relational operators are: union, difference, product, project, restrict, intersection, divide, and join. They can be used to extract information from the built database (Worboys, Duckham, 2004).

Relational models are not the most suitable ones for managing spatial data. (Worboys, Duckham, 2004) lists some of the main difficulties:

1) spatial data structure does not naturally fit with tabular structures (in many cases, the attributes should contain more than one value, violating the first normal form, which requires atomic values); for example, polylines are defined by a set of points, which are in turn composed by three coordinates, and it is impossible to store this information in a single attribute;
2) the required performance is high, and common computational models can slow down the system;
3) spatial data require specialized indexing, since the investigations can be performed through a single attribute value, but also, for example a point or a range of values in an area, that is more difficult than simple linear indexing; extensive studies are available in the literature about this topic (Worboys, Duckham, 2004; Laurini, Thompson, 1992).

For solving some of these problems, some extensions to the basic relational model are possible: for example, it is possible to define personal data-types, operations, indexes or access methods and active database functions (e.g. triggers) (Worboys, Duckham, 2004). Available software permit the exploitation of these possibilities (for example PostgreSQL - PostGIS implements these functions).

2.2.2.4.1 Structured query language (SQL)

SQL is a language that permits both the definition of the database schema (data definition), and the manipulation of the data in the schema, from the filling in to the retrieving (Worboys, Duckham, 2004).

It was based upon relational algebra and tuple relational calculus; SQL consists of a Data Definition Language (DDL), data manipulation language, and a data control language. The scope of SQL includes data insert, query, update and delete, schema creation and modification, and data access control. Although SQL is often described as, and to a great extent is, a declarative language, it also includes procedural elements.

SQL was initially developed at IBM by Donald D. Chamberlin and Raymond F. Boyce in the early 1970s. This version, initially called SEQUEL (Structured English QUEry Language), was designed to manipulate and retrieve data stored in IBM's original quasi-relational database management system. In the late 1970s, Relational Software, Inc. (now Oracle Corporation) saw the potential of the concepts described by Codd, Chamberlin, and Boyce, and developed their own SQL-based RDBMS (Relational database Management System). In June 1979, Relational Software, Inc. introduced the first commercially available implementation of SQL, Oracle V2 (Version2) for VAX computers.

*SQL became a standard of the American National Standards Institute (ANSI) in 1986, and of the International Organization for Standardization (ISO) in 1987 (*ISO/IEC 9075-1:2008: Information technology– Database languages– SQL– Part 1: Framework). *Since then, the standard has been revised to include a larger set of features. Despite the existence of such standards, though, most SQL code is not completely portable among different database systems without adjustments.*[36]

Some extensions of SQL exist in order to permit the storage and management of spatial data and to add some other useful constructs, borrowed from the object-oriented programming languages, as occurred, for example, in the implementation of PosgreSQL - PostGIS[37].

2.2.2.5 The object-oriented model

In the translation from higher-level languages (used in the first steps of the modelling process) to lower-level languages (used by the computational environment) some complexity of information can be lost. This problem is termed ‘impedance mismatch’ (Figure 29). In order to avoid or minimize this problem, the object-oriented approach can be used (Worboys, Duckham, 2004).

An object contains the static aspects of the data (as the attribute values of a tuple), which is the object’s state; moreover, it models the dynamic behaviour of the system, as a set of operations to be performed under some conditions. The definition of an object can be therefore expressed as:

$$\text{object} = \text{state} + \text{behaviour}.$$

Objects interact by sending messages to each other, activating particular behaviours. The objects with similar behaviours are organized into classes (Worboys, Duckham, 2004). If the object-oriented approach is applied to each modelling level, the impedance mismatch is reduced. Also the metaphorical power of such a model is to be considered, since, being a sort of object-oriented approach very similar to the human way to conceive the reality, its information organization and communication power is enhanced.

[36] The italic text cites https://en.wikipedia.org/wiki/SQL

[37] http://www.postgresql.org/

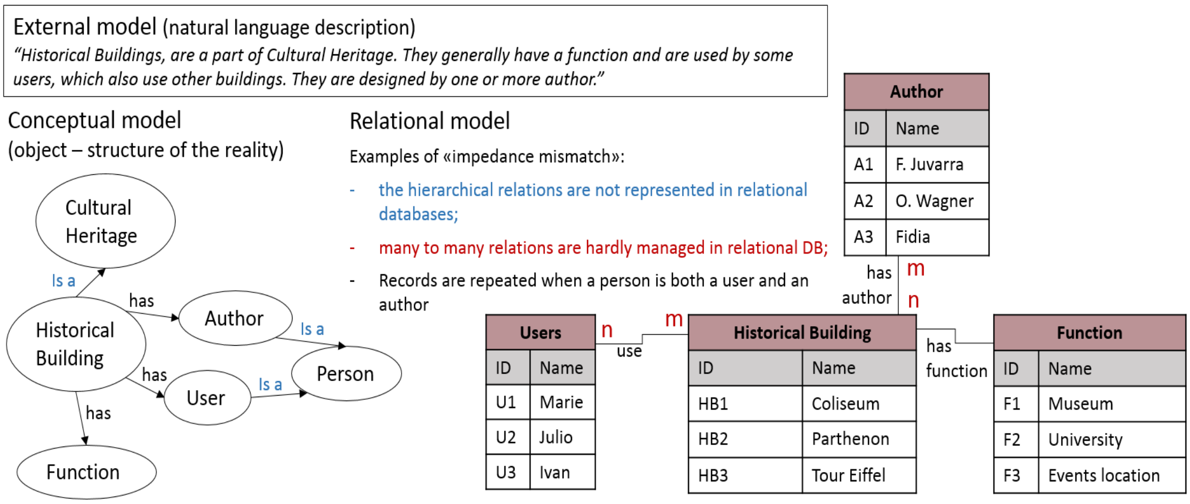


Figure 29. Example of the impedance mismatch problem: a definition in natural language (external model) a can be represented in a structured conceptual model (following the object-oriented model); however, in the translation towards a relational model, some complex characteristics are lost ('impedance mismatch').

Four constructs are important in object-orientation. Some of them are also included in extended relational databases (accordingly called 'object-relational' DBs) (Worboys, Duckham, 2004):

- *identity* - objects have unique identity, independent of attribute values. All the values can change but the identity remains unchanged;
- *encapsulation* - internal mechanisms of an object's behaviour are rigidly separated from the external access to, and effects of, that behaviour. It helps to reduce complexity by treating the behaviours an object exhibits separately from how an object achieves these behaviours. Relational DBs are based on the call-by-value principle, while values contained in an object are encapsulated and only accessible indirectly via an object's behaviour;
- *inheritance* - permits the establishment of hierarchies (with the inverse mechanisms, generalisations and specialisations), in which subclasses share properties of the superclass. It can be of a subclass towards a unique subclass (single inheritance) or towards more than one superclass (multiple inheritance);
- *polymorphism* - the ability of an object to fulfil different roles in different contexts, due to the inheritance relationships. An instance of a subclass can be substituted by an instance of a superclass (this is a particular form of polymorphism, called 'inclusion polymorphism'). Another type of polymorphism is the implementation by a subclass of its own specialized algorithms for achieving generalized behaviour (overloading).

A further component of the object-oriented structure is the 'association': it groups objects together, in order to model phenomena with complex internal structure. A special type of association, very common and useful in spatial application modelling, is aggregation, which concerns the part/whole association between objects (the 'part-of' relation). It is an asymmetric relation and can be organized into a hierarchy of aggregation relationships, which is termed 'partonomy' or 'mereological hierarchy'. An association can be 'homogeneous' (all the associated objects belong to the same class) and 'ordered' (if the sets of objects follow a specific order).

The complexity of the object-oriented approach causes the performance of object- oriented DBs to be lower than that of relational DBs, together with some problems of query optimisation and other implementation issues (Worboys, Duckham, 2004). Furthermore, it is not the best choice for dealing with surfaces spatial gradients (objects not to be defined as an object) (Laurini, Thompson, 1992). This problem is also identified by Galton (2005), who discusses the necessity to develop a formal ontological framework for representing the relationship between objects and fields describing unbroken continua. He calls that ‘multi-aspect phenomena’. However, its representation potentialities are higher, and recent standard languages for modelling some application fields (including spatial ones) use object orientation for defining the standard data models.

Object orientation demands an *a priori* definition of the objects and is therefore seen by some authors (e.g. Laurini, Thomson 1992) as more intuitively appealing but practically more difficult.

2.2.2.5.1 Object-Oriented DataBase Management System (OODBMS)

Some OODBMSs exist (e.g. EyeDB www.eyedb.org), but they are not specifically designed to manage spatial data. Most DBMS and GIS management software systems follow the relational model, so that they cannot manage the useful characteristics of object-oriented (OO) systems. Some Object-Oriented GIS (OOGIS) were implemented in the past in order to manage important characteristics of the semantic structure of spatial data (Mennis, 2003) (e.g. the project GODOT, Gaede, et al., 1994; or O2, Scholl, Voisard, 1992; Gong, Li, 1996; Egenofer, Frank, 1992). However, today the most widespread GIS management software packages (whether commercial or open-source) do not offer these functions. The more advanced systems use a hybrid object-relational model (ORDBMS), that is, a relational model that includes some functionalities of the OODBMS, such as inheritance and polymorphism.

The scarce success OODBs had at the time of their introduction is explained by (Bloor, 2003) by the fact that by the time object databases were introduced, relational database vendors had already gained momentum and market penetration. It was easy to write OO routines to access relational databases with their standards-based SQL interfaces. In contrast, most early object databases provided no SQL capability at all and were not suitable for query applications. Therefore, the database market never devoted a lot of attention to object databases until the emergence of XML (eXtensible Markup Language) (Section 7.4.1) as a data definition language. Recently the object-oriented paradigm was taken into account again for modelling spatial DBs (Worboys et al., 2006) and became essential to use defined XML-based standards.

Ontologies exploit object-oriented structures and systems, which are unusual in current and known GIS management systems: some of the most widespread ones (*PostgreSQL-PostGIS*, *ArcGIS Geodatabase tool, Oracle*) implement object-relational systems, which are hybrid systems including some constructs of object-oriented databases, but not the whole potentiality. Some studies about the development of some semantic GISs have been performed, beginning in the mid-1990s (Mennis, 2003; Fonseca et al., 2002). In these studies, an object-oriented approach was used as an effective solution for expressing and storing data meanings (Scholl, 1992). In this way, even more powerful systems could be built with significant data interoperability and a reduction of any potential ambiguity. However, at the present moment, few similar systems are available, with users preferring to employ SQL-based implementations (Belussi et al., 2011). This is due to the necessity to adapt the exigencies to the available platforms and software systems and to the necessity to change the storage methods to permit the production and management of computationally heavy files. In the coming years, probably the object-oriented GIS will be developed again or, some different interfaces from the GIS we know will be improved to include spatial analysis and query functionalities in an OO approach.

Sheet 1. The object-oriented database system manifesto

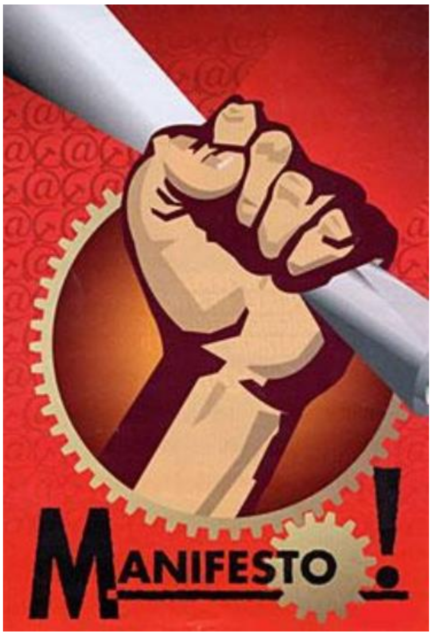


Figure 30. Image about the Object-oriented database system manifesto[38]

The following text is part of the document (Atkinson et al., 1989) and represents a synthesis of its content for what is of interest in this thesis.

This paper attempts to define an object-oriented database system. It describes the main features and characteristics that a system must have to qualify as an object-oriented database system.

We have separated the characteristics of object-oriented database systems into three categories: mandatory (the ones that the system must satisfy to deserve the label), optional (the ones that can be added to make the system better but which are not mandatory) and open (the places where the designer can select from a number of equally acceptable solutions). [39]

Mandatory features: the Golden Rules

Complex objects: *Complex objects are built from simpler ones by applying constructors to them.*

- ***Object identity:*** *in a model with object identity, an object has an existence which is independent of its value.*

- ***Encapsulation:*** *(i) the need to cleanly distinguish between the specification and the implementation of an operation and (ii) the need for modularity. The idea of encapsulation in programming languages comes from abstract data types. In this view, an object has an interface part and an implementation part. The interface part is the specification of the set of operations that can be performed on the object. It is the only visible part of the object. The implementation part has a data part and a procedural part. The data part is the representation or state of the object and the procedure part describes, in some programming language, the implementation of each operation. The database translation of the principle is that an object encapsulates both program and data. We*

[38] http://www.slideshare.net/signer/the-object-oriented-database-system-manifesto

[39] The italic text cites (Atkinson et al., 1989, The Object-oriented database system manifesto).

believe that proper encapsulation is obtained when only the operations are visible and the data and the implementation of the operations are hidden in the objects.

- ***Types and Classes**: there are two main categories of object-oriented systems, those supporting the notion of class and those supporting the notion of type.*
    - *A **type**, in an object-oriented system, summarizes the common features of a set of objects with the same characteristics. It corresponds to the notion of an abstract data type. It has two parts: the interface and the implementation (or implementations). Only the interface part is visible to the users of the type, the implementation of the object is seen only by the type designer. The interface consists of a list of operations together with their signatures (i.e., the type of the input parameters and the type of the result). The type implementation consists of a data part and an operation part. In the data part, one describes the internal structure of the object's data.*
    - *The notion of **class** is different from that of type. Its specification is the same as that of a type, but it is more of a run-time notion. It contains two aspects: an object factory and an object warehouse. The object factory can be used to create new objects, by performing the operation new on the class, or by cloning some prototype object representative of the class. The object warehouse means that attached to the class is its extension, i.e., the set of objects that are instances of the class.*
- ***Class or Type Hierarchies:** Inheritance has two advantages: it is a powerful modeling tool, because it gives a concise and precise description of the world and it helps in factoring out shared specifications and implementations in applications. There are at least four types of inheritance:*
    - *In substitution inheritance, we say that a type t inherits from a type t', if we can perform more operations on objects of type t than on object of type t'. Thus, any place where we can have an object of type t', we can substitute for it an object of type t. This kind of inheritance is based on behavior and not on values.*
    - *Inclusion inheritance corresponds to the notion of classification. It states that t is subtype of t', if every object of type t is also an object of type t'. This type of inheritance is based on structure and not on operations.*
    - *Constraint inheritance is a subcase of inclusion inheritance. A type t is a subtype of a type t', if it consists of all objects of type t which satisfy a given constraint. An example of such a inheritance is that teenager is a subclass of person.*
    - *With specialization inheritance, a type t is a subtype of a type t', if objects of type t are objects of type t which contains more specific information. Examples of such are persons and employees where the information on employees is that of persons together with some extra fields.*
- ***Overriding, overloading and late binding:** In an object-oriented system, we define the display operation at the object type level (the most general type in the system).*
- ***Computational completeness:** it simply means that one can express any computable function, using the DML () of the database system.*
- ***Extensibility:** there is a means to define new types and there is no distinction in usage between system defined and user defined types.*
- ***Persistence:** is the ability of the programmer to have her/his data survive the execution of a process, in order to eventually reuse it in another process.*

- ***Secondary storage management:*** *It is usually supported through a set of mechanisms. These include index management, data clustering, data buffering, access path selection and query optimization.*
- ***Concurrency:*** *the system should offer the same level of service as current database systems provide.*
- ***Recovery:*** *in case of hardware or software failures, the system should recover.*
- ***Ad Hoc Query Facility:*** *to provide the functionality of an ad hoc query language. For instance, a graphical browser could be sufficient to fulfil this functionality.*

Optional features: the goodies

- *We put under this heading things which clearly improve the system,*
- *Multiple inheritance*
- *Type checking and type inferencing*
- *Distribution*
- *Design transactions*
- *Versions*

Open choices

- *Programming paradigm*
- *Representation system*
- *Type system*
- *Uniformity*[40]

[40] The italic text cites (Atkinson et al., 1989 The Object-oriented database system manifesto).

2.2.3 An emerging alternative: BIM and HBIM (Historical BIM)

A Building Information Model (BIM) simulates a construction project in a virtual environment. It was developed in the early 2000s to serve as a support for project activities. It is a 3D computer-generated model containing geometry and relevant data needed to support the construction, fabrication and procurement activities required to realize a building. From the model, views and data appropriate to various users' needs can be extracted and analysed to generate information that can be used to make decisions and to improve the process of delivering the facility (Azhar et al., 2008).

It is structured as an object-oriented archive integrating geometric data deriving from parametric modelling and related data, which semantically enrich each element. Three aspects are involved: the geometrical model of the component, the attribution of categories and material properties to the components and the relations between them (Hichri et al., 2013).

In BIM, it is also possible to store some information regarding topology and directional spatial relations among the objects.

A building information model can be used for the following purposes (Azhar et al., 2008):

- *visualization*: 3D renderings can be easily generated in-house with little additional effort;
- *fabrication/shop drawings*: it is easy to generate shop drawings for various building systems;
- *code reviews:* fire departments and other officials may use these models for the review of building projects;
- *forensic analysis*: a BIM can be adapted to graphically illustrate potential failures, leaks, evacuation, etc.;
- *facilities management*: facilities management departments can use BIM for renovations, space planning, and maintenance operations;
- *cost estimating:* BIM software has built-in cost estimating features. Material quantities are automatically extracted and changed when any changes are made in the model;
- *construction sequencing:* a building information model can be effectively used to create material ordering, fabrication, and delivery schedules for all building components;
- *conflict, interference, and collision detection*: because BIM models are created, to scale, in 3D space, all major systems can be visually checked for interferences. This process can verify that piping does not intersect with steel beams, ducts, or walls.

A special category of these models represents not a project of a building being built but existing surveyed buildings. These are named 'as-built BIM' or 'as-is BIM'. These would inform about the state of conservation of historical buildings. Different approaches exist for detecting sub-parts of a building being represented in such BIMs; they are summarized in (Hichry et al., 2013).

For geometric data, often B-Rep are used, usually derived from a parametric modelling activity. For existing buildings, in particular historical architecture, some library of parametric objects based on historical data exist, called 'Historical BIM' (HBIM). Moreover, some studies for extracting profiles for mouldings were performed (Hichry et al., 2013).

Parametric modelling, even if it constitutes an advantage in many cases for understanding the building, for projecting new parts, or for developing hypotheses related to nonexistent (destroyed or never realized) building parts' shape, it is still too rigid to adequately represent the irregular geometries that often characterize historical buildings. For this reason, in this research, the triangulated mesh approach is preferred, in order to avoid losing surveyed shape details in approximations.

On the other hand, the segmentation possibilities offered by software managing BIM and the possibility to include other information about the spatial relationship between objects make a possible integration between GIS approaches (which support the management and analysis of information perhaps more tailored to surface definition and description) and BIM a challenging issue.

In a similar way to GIS, BIM also has a standard structuring data model, the Industry Foundation Classes (IFC). It is intended to describe building and construction industry data. It was developed by Autodesk (1994) and is now maintained by buildingSMART (formerly the International Alliance for Interoperability, IAI) to facilitate interoperability in the architecture, engineering and construction (AEC) industry. The IFC model specification is registered by ISO and is an official International Standard ISO 16739:2013[41].

Usually CityGML, which is a standard data model for cartography, further described in Section 6.3.4, and IFC are considered as structuring different levels of detail (one for urban objects and the other for small building details). This point of view is not here shared, since CityGML includes details that go beyond the urban level, foreseeing different levels of detail from the regional scale to the building. This is the reason why in this research it was chosen for being extended. It is anyway undeniable that the two classifications were developed for distinct aims: the project of the new (starting from the detail of building installations and construction details) or the cartographic representation (from the city level to reach the interiors of the buildings). Given that an integration is probably the best solution, some studies have worked in this sense, by proposing the mapping of the two standards to an intermediate model (El-Mekawy et al., 2012; El-Mekawy, Östman, 2010). Others have instead used the possibility to extend the CityGML model for processing a GeoBIM Application Domain Extension (ADE, mechanism for extending CityGML), which extends the CityGML model with the IFC entities (de Laat, van Berlo, 2011). This is an interesting approach. In this thesis a parallel approach was chosen, for permitting a better flexibility in the definition of the geometries and in the semantic labelling of them, for which the IFC vocabulary, born to represent new buildings is not sufficient. An integration between the models should however be interesting, and it is possible since the same extension method and structures (the CityGML ADE) were used.

[41] https://en.wikipedia.org/wiki/Industry_Foundation_Classes

2.3 The analysis

2.3.1 GIS analysis of architectural heritage data

A number of analyses can be performed in GIS. The more usual ones are overlay operations, buffering, network analysis and similar others. Others consider thematic data for performing statistics, interpolations, graphs and thematic visualizations or query (Figure 31). All of them are affirmed possibilities of traditional 2D GIS.

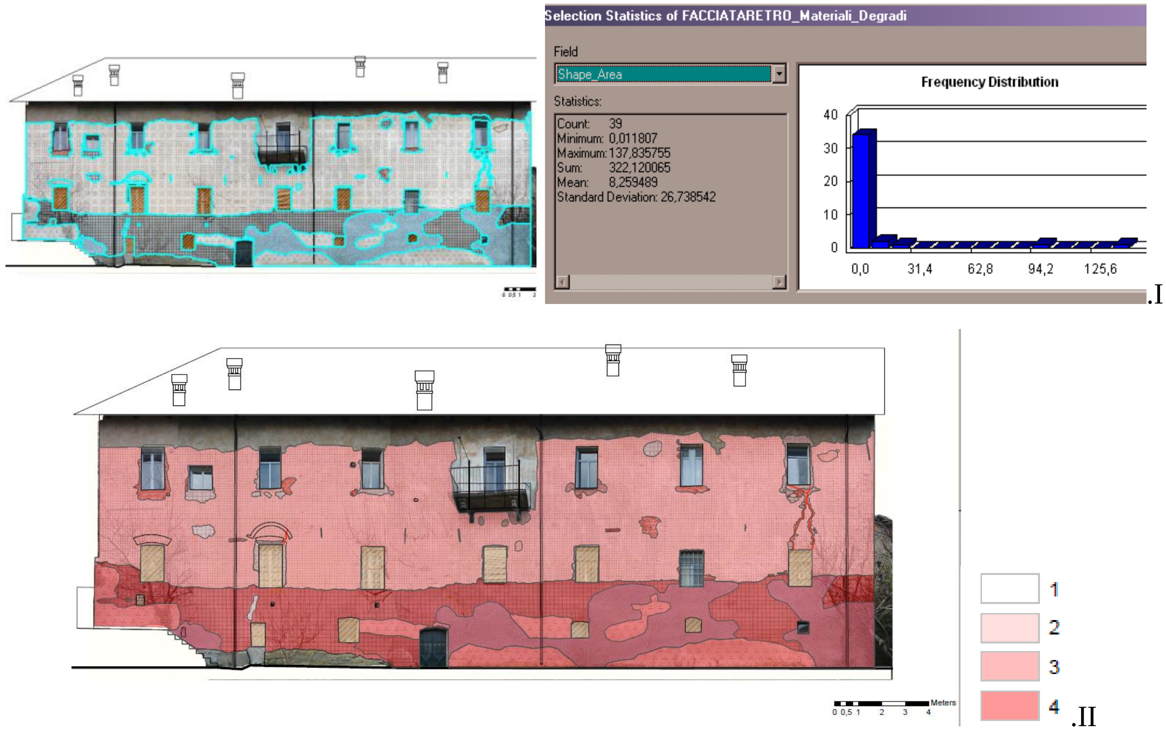


Figure 31. Examples of analysis of thematic data stored in a database (visualized through the geometries with which they are associated). In (I) a query is made for selecting a kind of surface (having the similar pathology) and statistics are computed. In (II) some parameters stored as thematic values and overlay characters are considered to produce a thematic representation of the level of urgency of the intervention on the façade.

Another family of analysis that can be performed in GIS is spatial analysis. These consider mainly 2.5D geometric data (such as DTM, DSM, and triangulated terrain models) for extracting information about their shape and how this can act in relation to some phenomena. Since these tools were developed for environmental management goals, the surface analysis is aimed to understand the behaviour of the surface conceived as the terrain.

Some kinds of surface analysis are:

- the ones that analyse the surface itself (e.g. aspect, slope mapping); these have influence on other algorithms, for example the ones computing the rate of solar radiation received by the surface;
- others estimate the water itinerary on the surface, providing a set of products useful for hydrological studies (basins, watershed, deepest path, etc.);

- a further category can be used to estimate the viewshed from specific points or lines.

The same processing can be effectively borrowed by the architectural heritage application field, for which it can be a powerful tool.

Some examples follow about the application of the processing of the first listed type (Figure 32) and the second type (Figure 33 - Figure 34) to large-scale elevation models.

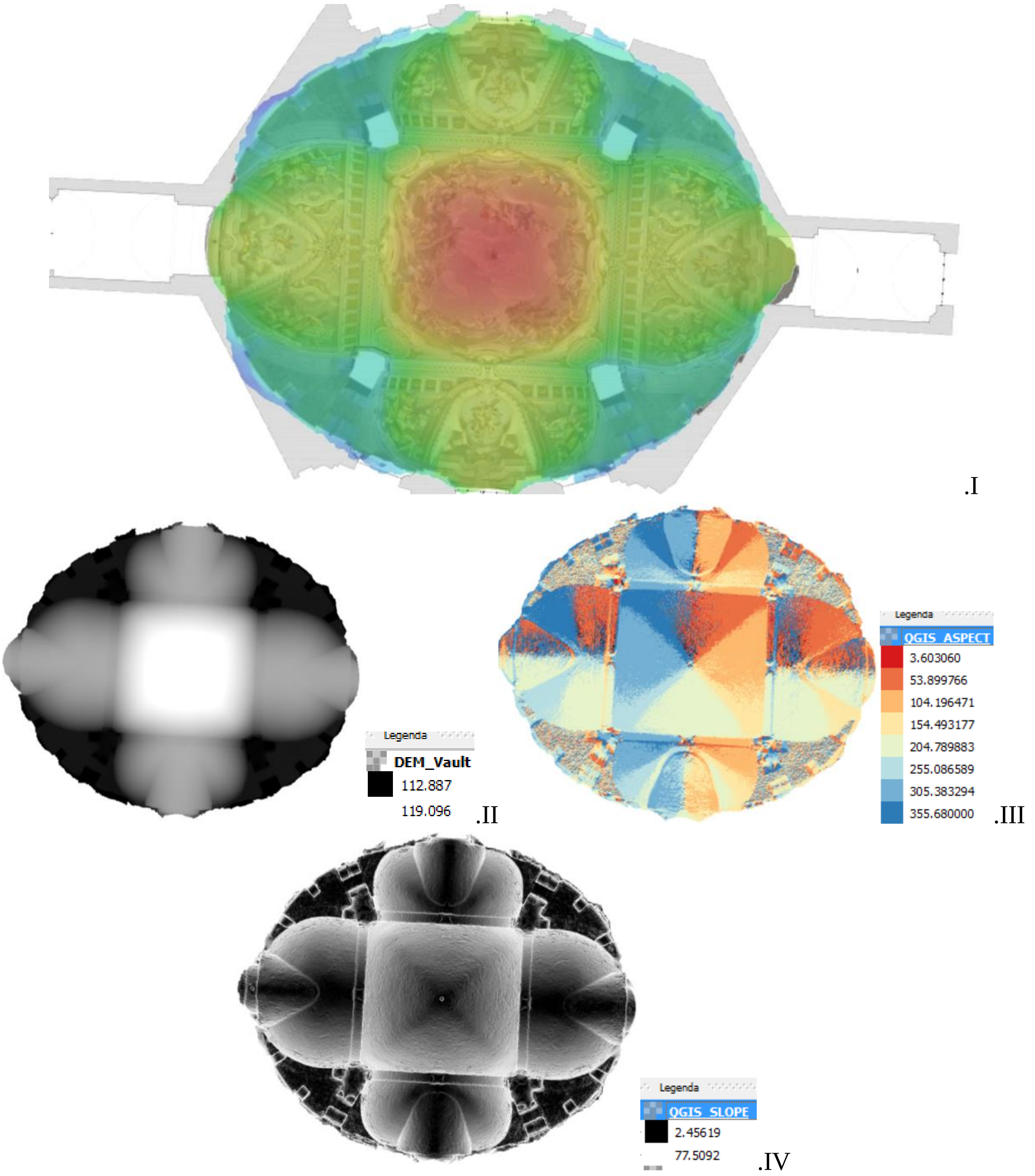


Figure 32. Automatic aspect (III) and slope (IV) extraction by the DEM (II) of the vault of the Hall of Honour of Stupinigi royal residence (I: 2D plan of the hall, with orthophoto of the vaults and DEM superimposed), generated using Structure from Motion image processing. The GIS spatial analysis (in QGIS[42]) enlighten possible deformation characteristics.

[42] http://www.qgis.org/it/site/

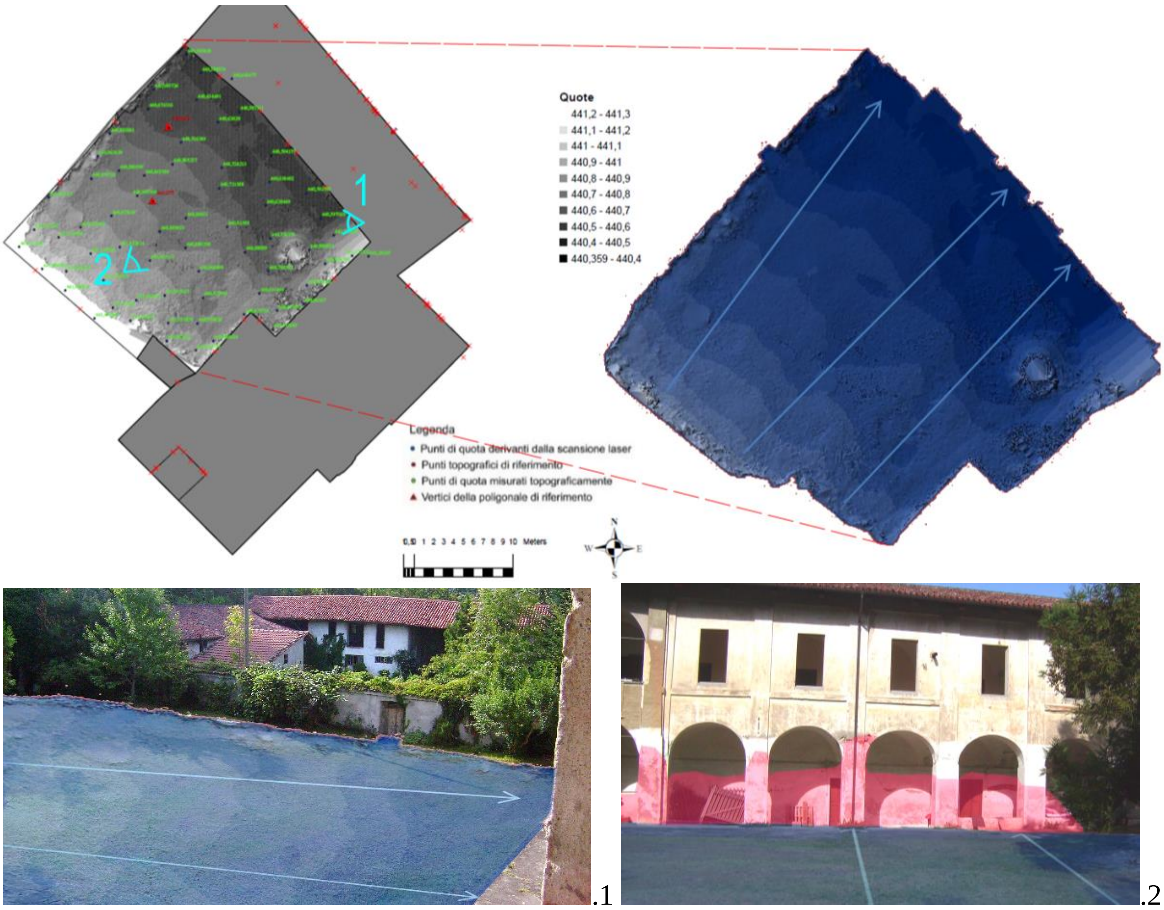


Figure 33. The DTM of the court of a convent (from a TLS survey) used for hydrological analysis: the north-east building suffers humidity problems (highlighted in red in image n.2), so the direction of the water on the court was investigated (Noardo, 2012).

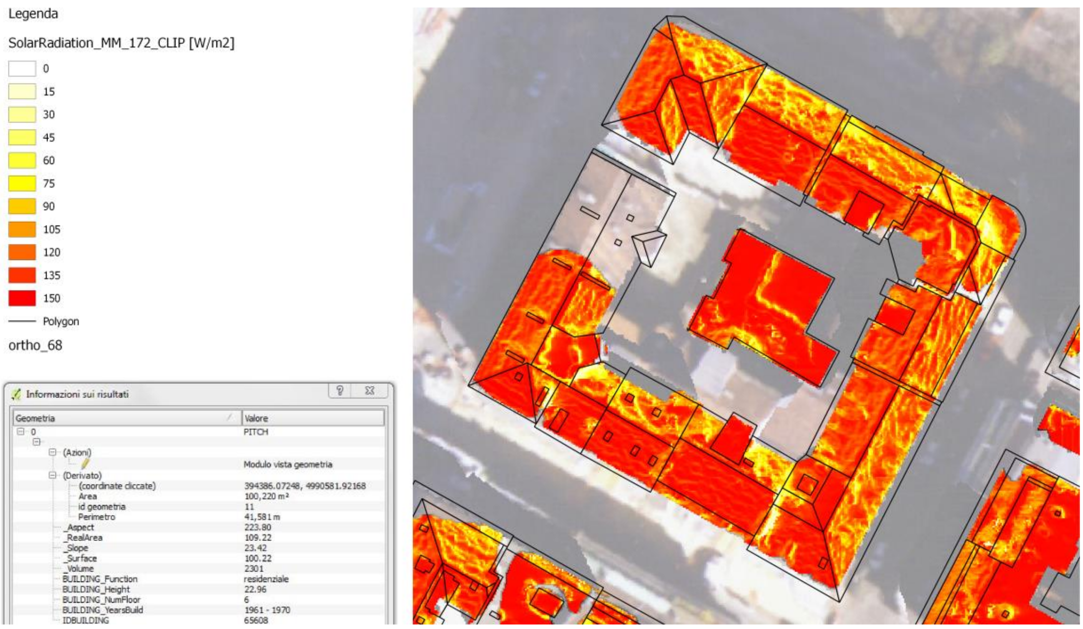


Figure 34. Example of solar radiation estimation in QGIS on the roofs of some buildings in the centre of Turin (Italy). It can be useful as a support for energetic policies or for investigating possibly connected pathologies of the buildings (Chiabrando et al., 2016).

2.3.2 3D models as a support for data analysis

When geometry is more complex than for a 2.5D representation, different supports are needed, such as the proper 3D models. Algorithms are in development for extracting information (such as profiles, contours, or deformation mapping) directly from the 3D meshes (Figure 35).

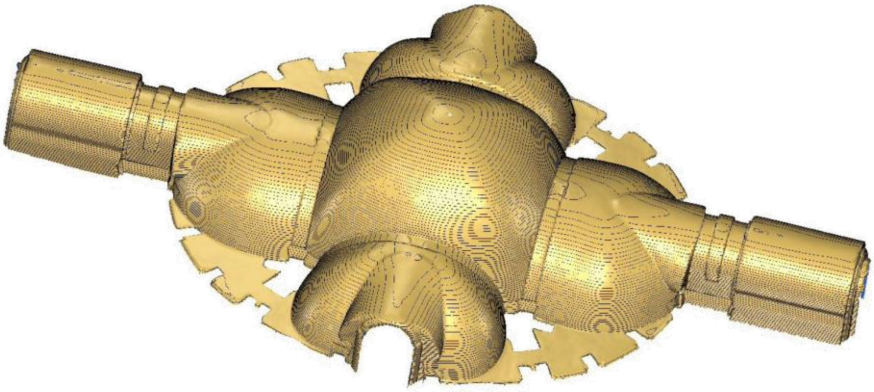


Figure 35. Contour lines extracted on the 3D model (model processed by E. Donadio)

Other kinds of analysis consist in generating geometry reconstruction for supporting the historical investigation and for verifying the employed building techniques (Figure 36 - Figure 37).

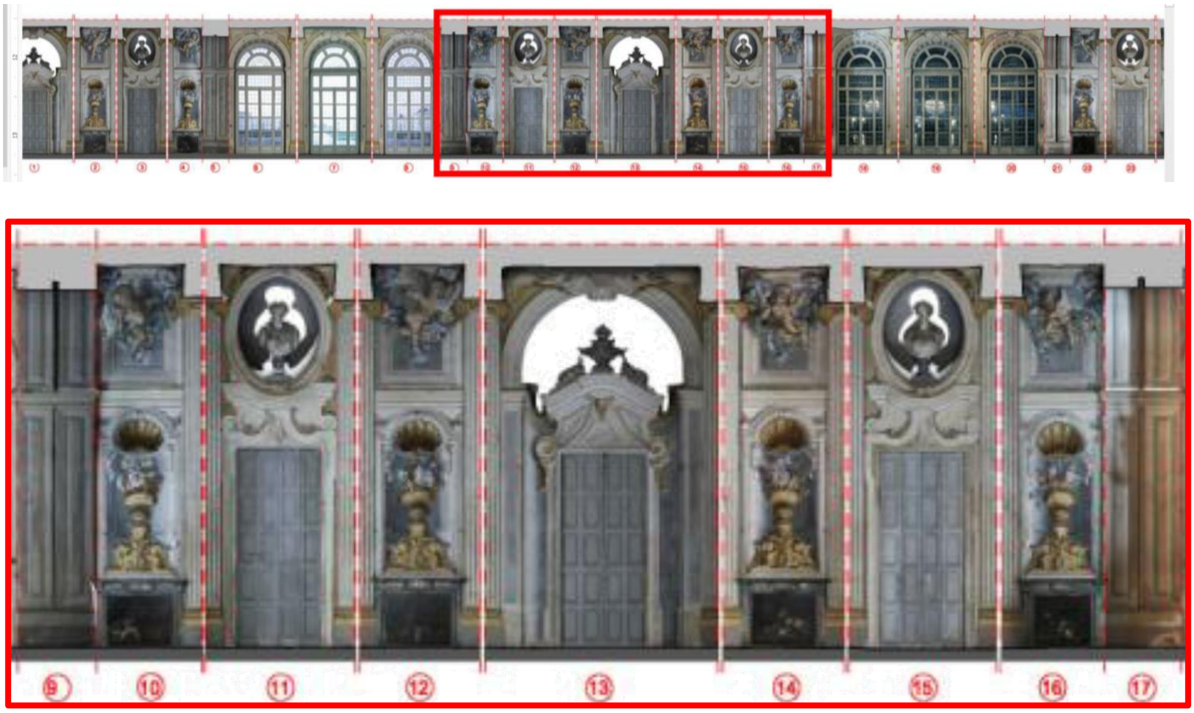


Figure 36. Unrolling of the first order in the interior façade of the hall of honour in the Stupinigi royal redidence (the whole development in the first image and the central part in the second one). From this product (obtained by the 3D model) some composition rules can be studied (processing by E. Donadio).

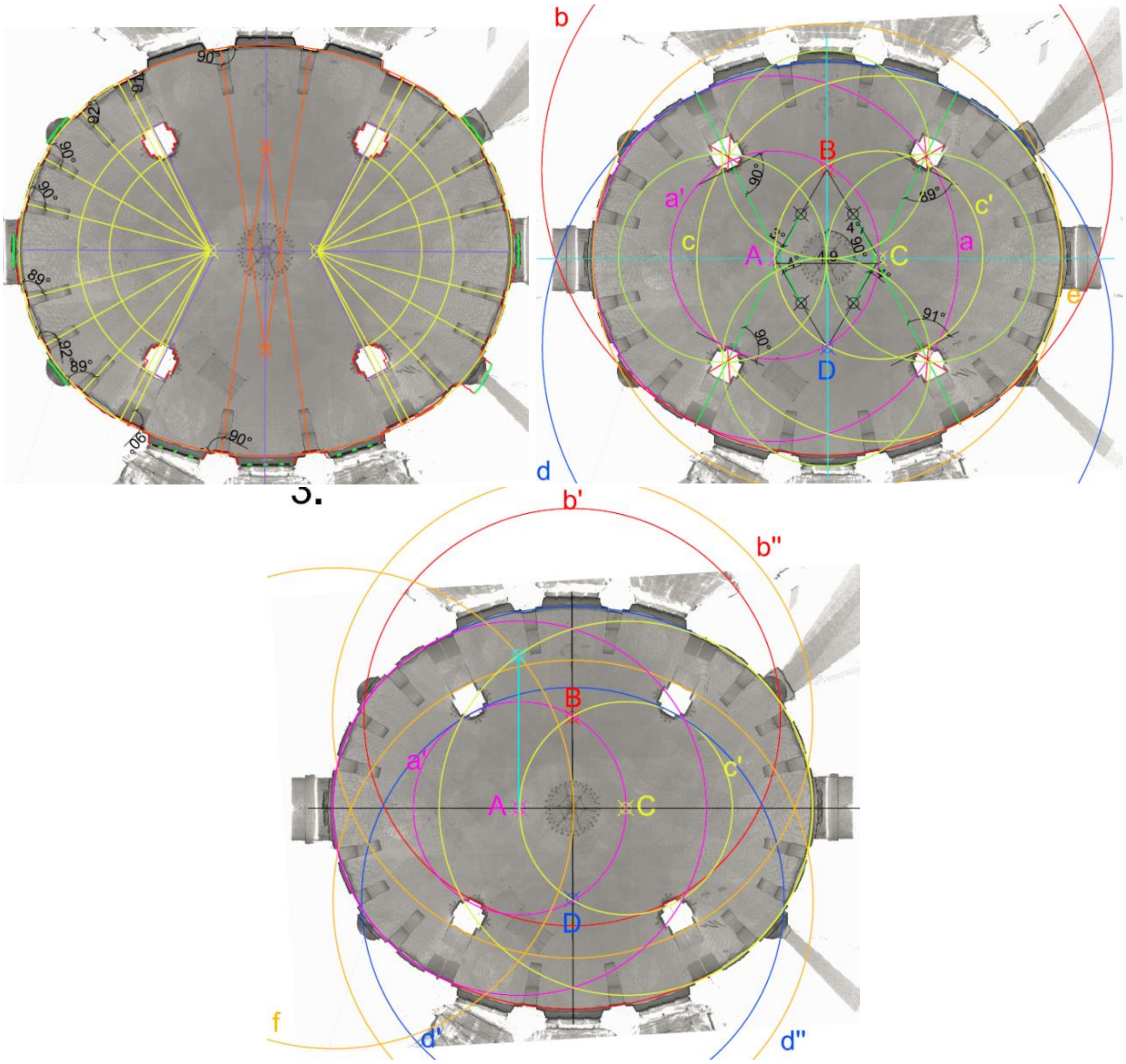


Figure 37. Some images of the study of the proportions and geometry of the Stupinigi hall of honour, on the base of the laser scanner survey.

Moreover, studies are being developed for using the models directly, for example, for static behaviour analysis of the objects (Chiabrando et al., 2015).

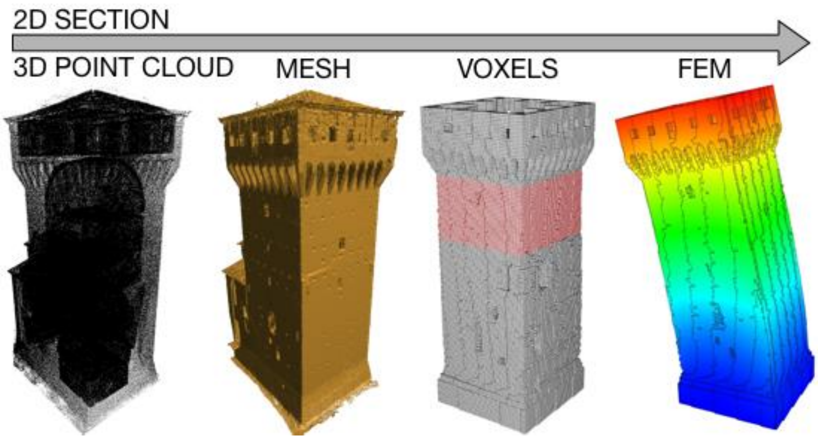


Figure 38. Example of FEM (Finite Element Model) analysis of a tower surveyed with TLS (Castellazzi et al., 2015).

2.4 The communication

The obtained archives can be effectively used to communicate to users, researchers, or various stakeholders the information or specific results. The formats of communication products can vary following the requirements of the users, the nature of the built system (2D-3D, etc.), and the chosen communication channels (paper maps, web maps, digital products).

A particular type of communication product, which is spreading for its great potentialities and which is also among the aims of this thesis, is the 3D digital model, possibly enriched with semantic values. For the use of such kind of products, the London Charter for the computer-based visualisation of cultural heritage (2009) has stated some principles. In its document, a set of rules regard the documentation; among other aspects, the formats and standards aspects of documentation are considered.

Documentation Formats and Standards

4.11 Documentation should be disseminated using the most effective available media, including graphical, textual, video, audio, numerical or combinations of the above.

4.12 Documentation should be disseminated sustainably with reference to relevant standards and ontologies according to best practice in relevant communities of practice and in such a way that facilitates its inclusion in relevant citation indexes.[43]

2.4.1 The traditional channels and the support of GIS in the generation of information products

GIS can generate layouts similar to traditional paper maps and drawings. They are effective for operators used to having these kinds of products as a support for their activities. In this form, the part of information of interest implemented in the GIS is visualized in a static product, which can be assimilated to traditional supports for analysis and projects. This is sometimes indispensable for communicating with non-expert users of these systems. The advantage of GIS, also in this case, is the possibility to produce a quantity of different layouts from the same representation. Also traditional graphs and table reports can be exported if needed. Some examples of similar layouts are represented in Figure 39 - Figure 40.

Figure 39. Mapping of the pathologies of a convent front, as a support for preservation intervention planning (Noardo, 2012).

[43]The italic text cites the London Charter for the computer-based visualisation of cultural heritage.

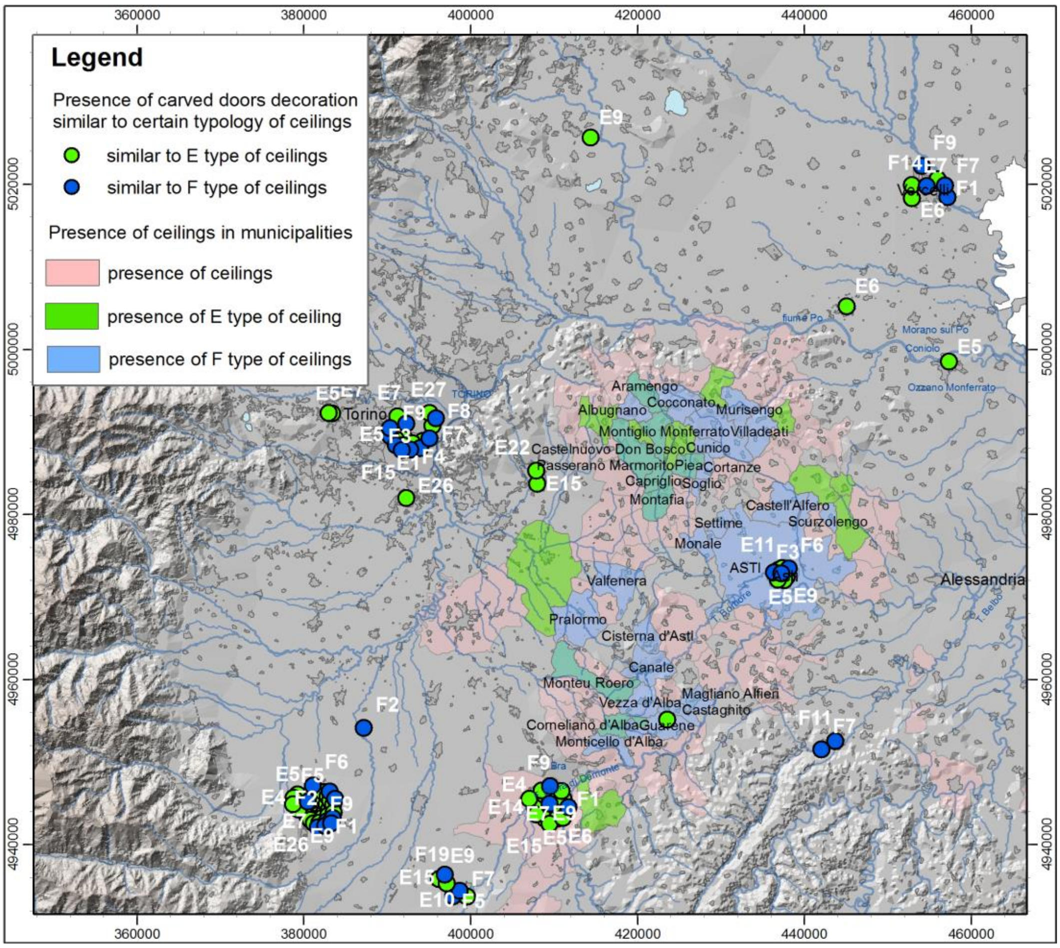


Figure 40. Map of the presence of a particular kind of cultural heritage, the gypsum ceilings in a Piedmont area (Italy) (Spanò, 2013).

2.4.2 WebGIS

An effective solution for sharing the produced information is the publishing of webGIS, that is, a GIS published on the web for the information to be visualised, queried, exported, or analysed. Express Open Geospatial Consortium (OGC) web services are available: Web Map Service (WMS), Web Feature Service (WFS), Web Coverage Service (WCS) for data export, or Web Processing Service (WPS) for processing and analysis. A list of these and further similar OGC standards that enable these possibilities are listed in http://www.opengeospatial.org/standards/.

The publishing of the data can exploit some web platforms, that are constituted by servers, storage devices, and communication procedures, permitting the sharing of GIS structured data with connected meta-information, visualisation styles, hyperlinks, and accessibility permissions.

For these technologies, several architectures exist for different requirements (e.g. Pascaul et al., 2012; Milosavljević et al., 2005). However, one of the most important open source platforms for publishing webGIS is GeoServer (geoserver.org). It is an OGC-compliant implementation to be used as a server for sharing (possibly open) geospatial data with aims of interoperability.

A further interesting open source platform, which in turn exploits GeoServer, for building WebGIS and spatial data infrastructures (SDI) is GeoNode[44] (Agosto et al., 2011).

The WebGIS aims can be various: from the simple visualisation of the data, to the base for interoperability and public-participatory activities, for simple data retrieving or as a support for decision-making (Gallo, Malatacca, 2012; Boroushaki, Malczewski, 2010; Miler et al., 2010; Dragićević, Balram, 2004) or even emergency management[45].

Cultural heritage WebGIS or similar web-platforms exist. One of them is the Italian 'Risk map' (Section 6.2.6); others are produced to share research data or to provide support for various levels of users or operators (Figure 41). Some of them are briefly described in the examples of GIS for architectural heritage management (Chapter 3).

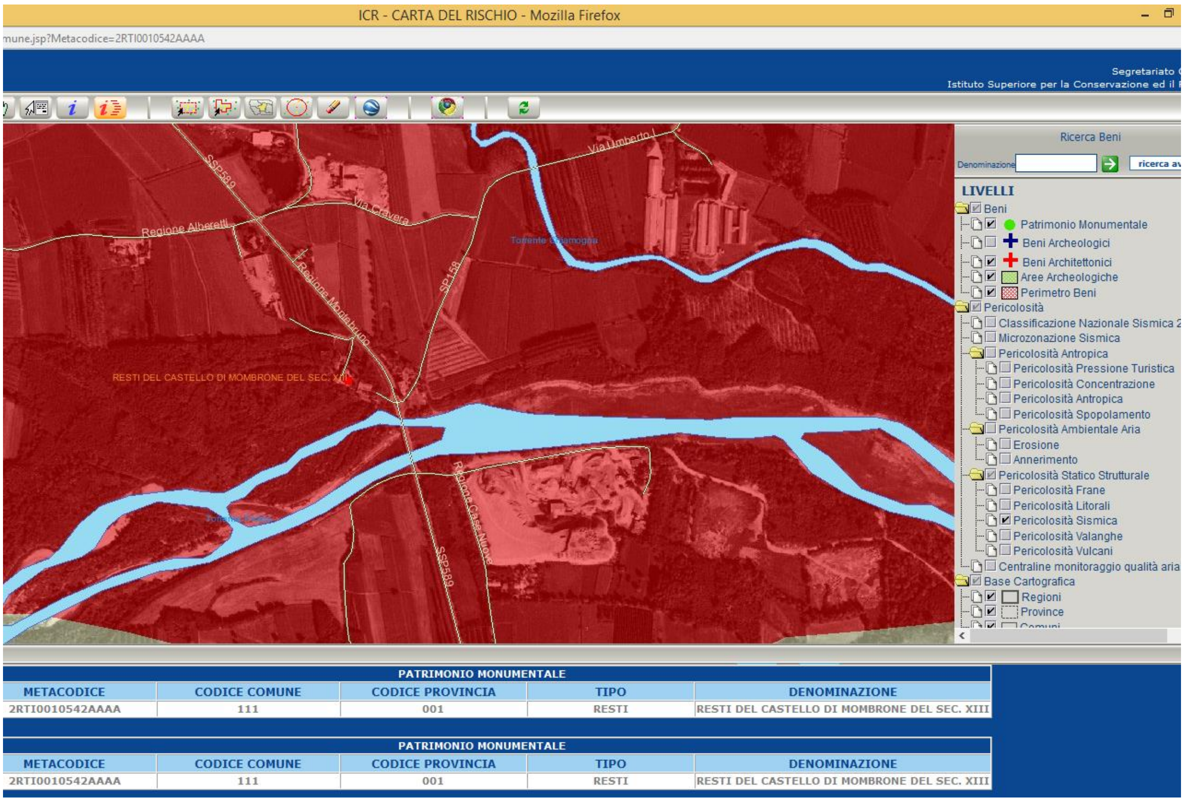


Figure 41. Example of research on the webGIS of the Italian Risk Map: a protected cultural item is mapped and related to the possible risks (in the image, red represents the seismic risk)[46].

[44] geonode.org/
[45] http://www.agire.it/
[46] http://www.cartadelrischio.it/utenti/home_utenti.asp#

3 GIS FOR LARGE-SCALE ARCHITECTURE INFORMATION: STATE OF THE ART

"Cultural heritage resource management is a cyclical process of caring for cultural heritage resources. The process comprises the following activities: documentation, evaluation, strategy development, intervention, monitoring and strategy review. GIS can be used to assist all activities in the process."[47]

The advantages of GIS for Cultural Heritage (CH) management consist in the possibility given by the database component to extract data, in the typical data retrieving and querying the information; spatial analysis and statistics computations are further function of these instruments. The automatic thematic maps also offer significant advantages for the communication of results or of selected information.

Some limits persist in the difficulty to manage 3D data by means of common GIS software: also when a 3D visualisation is possible (e.g. in ESRI ArcScene), the functions that make a GIS a unique instrument are weaker. For example, editing possibilities are reduced, spatial processing is quite impossible, the application of symbols to the objects is not always supported, and also when the functions are possible, the hardware computing performances are often insufficient to execute them. For the 3D management of data, anyway, several tools and applications are being programmed, especially for sharing and retrieving the information on the web.

A further limiting aspect is the relational model, which structures most of the common GIS software. Even if some DBMSs can implement some improved functions and constructs typical of the object-oriented approach, some complexity requirements expressed by the structuring models based on standards can be lost. The relational model is, however, best suited for other kinds of data storing, for example the long dynamic series of monitoring data. Some expressly programmed (often web-based) platforms are moving towards more object-oriented applications.

In this chapter, some examples of the use of GIS for cultural heritage are described. The technology spread in the field around the year 2000, experiencing a further improvement in the last years. A document (Petrescu, 2007) summarizes the spread of this technology in the world at the half-way point, the year 2007.

In the projects examined, the common and widespread GIS software tools are often used. In many cases the ESRI[48] software house products (ArcMap, ArView, ArcInfo, etc.) are used. Others employ open source software, such as QGIS[49]. The same is true for the DBMS: Oracle[50] products or other open source alternatives, such as PostgreSQL[51] are often cited. A summary and brief description of the available tools can be found in (Jafari Salim, 2014). Some projects, instead, programmed new platforms or applications either by using the foreseen languages, which permits the integration of the

[47] The italic text cites (Box, 1999)
[48] http://www.esri.com/
[49] http://www.qgis.org/en/site/
[50] http://www.oracle.com/index.html
[51] http://www.postgresql.org/

GIS existing software tools (e.g. the Python language[52] or ArcObject in the case of ESRI products[53]) or by building new tools in further programming languages (e.g. JAVA[54], JAVASCRIPT[55], PHP[56]).

The examples reported in the next subsections are grouped basing on the main aim of the built GIS and are ordered in an approximatly chronological order. The text mainly cites the text of the references.

3.1 GIS for general documentation aims

3.1.1 Applying the 3D GIS DILAS to archaeology and CH projects

The original aim of the 3D GIS project DILAS (Digital Landscape Server) was the efficient generation, management, and visualisation of large 3D landscape and city models **(**Wüst, et al., 2004). The result of the first project phase was an operational prototype of a 3D GIS based on an object-relational DBMS. The requirements include a strong support for object semantics, for complex object geometries and photo-realistic textures, for the integration of existing databases, and for temporal aspects.

3D object support in DILAS is based on an object-oriented topological 3D data model. This model is automatically mapped between an object representation in Java and an XML representation, which is stored in an object-relational DBMS (Oracle 9i). DILAS provides a rich semantic support at the object level and at the element level. DILAS is built on top of a spatial DBMS architecture (Oracle Spatial) and can make use of spatial indexing and querying functionality available in typical 2D GIS. This architecture has been extended to support complex 3D objects, e.g. buildings consisting of their exterior hull and of different types of interior objects (rooms, hallways, caverns, etc.) (Figure 42).

The goal for handling and manipulating 3D objects was to provide an optimum modelling flexibility in combination with an excellent object query and retrieval performance. The developed concept is based on the following components:

- a 3D object representation in Java and XML;
- a 3D object serialisation and de-serialisation;
- a persistence framework built on top of the DBMS;
- spatial data types for 3D and raster objects within an object-relational environment.

Many challenging functions are implemented (3D, semantics, topology). However, the employed data model is probably not standard, and the supporting software was expressly programmed, since little already existed; in fact, it is one of the major limitations to the use of similar systems.

---

[52] https://www.python.org/
[53] http://edndoc.esri.com/arcobjects/9.2/net/5bd93a2b-1c00-4927-ab26-5fbe3891a448.htm
[54] https://en.wikipedia.org/wiki/Java_%28programming_language%29
[55] http://www.w3schools.com/js/
[56] http://www.w3schools.com/php/

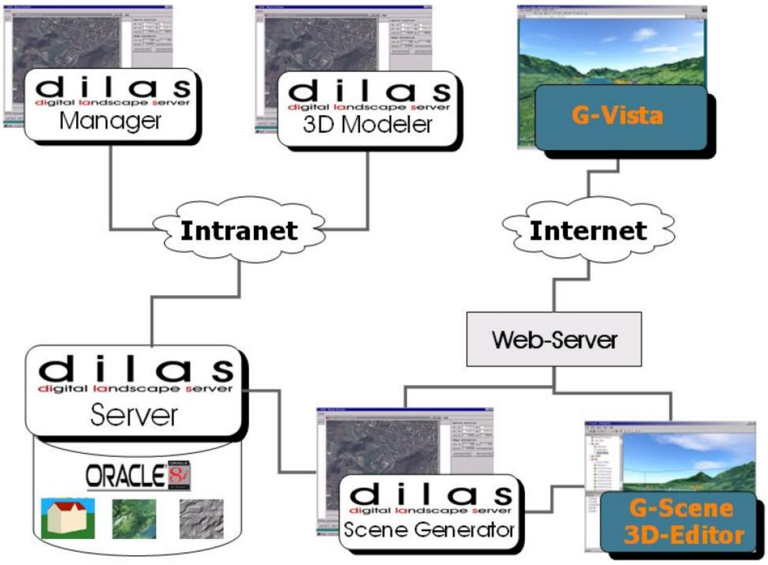


Figure 42. DILAS System Architecture (Wüst, et al., 2004).

3.1.2 SIARCH

The SIArch-Univaq, developed at the Architecture and Urbanism Department of L'Aquila University, is a part of a research study funded in 2006 by the Italian Ministry of the University and proposes the theme of an architectural GIS database conceived for the conservation and enhancement of architectural heritage (Centofanti, 2008). It is an information system developed at the architectural scale in order to make it easy to develop timely strategies for the conservation and enjoyment of architectural heritage. The database's structural organization was ultimately designed in order to make the available data integration possible together with the so-called 'Risk Map' national database (Section 6.2.6).

The aims are the possibility to query the model and to obtain and export 2D or 3D (Figure 43) thematic visualization (Figure 44) and useful synthesis or automatic computations and measurements as a support for a restoration process (including economic issues). Other issues are the storing of multimedia material (including raster and images) and further data useful for historical studies. ESRI ArcGIS was used in the development.

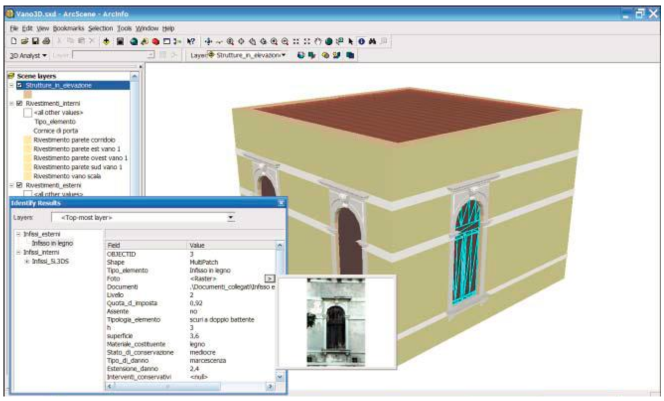


Figure 43. Example of 3D application of the system (Centofanti, 2008).

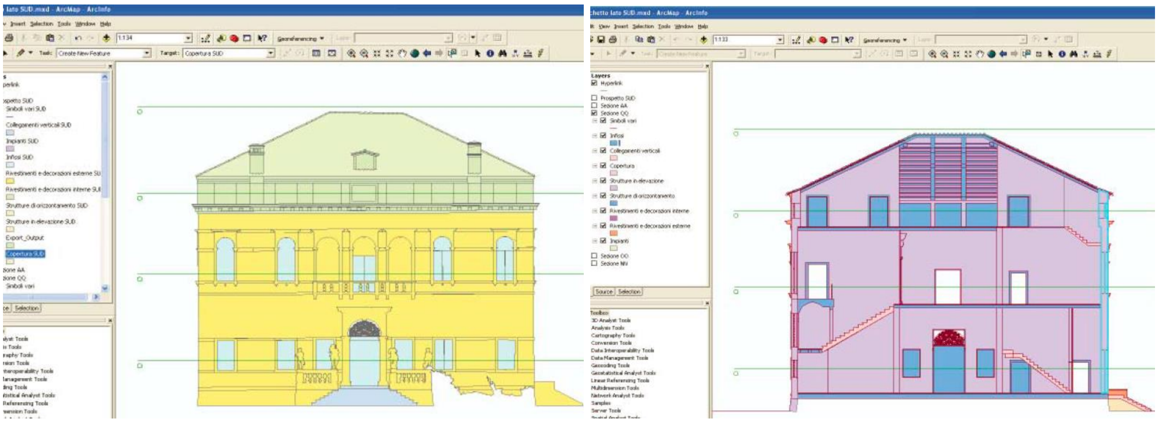


Figure 44. Examples of application of the system for the mapping of 2D vector drawings. The different projection plans of the objects (fronts, plans, sections) are interrelated through a hyperlink that is activated on the section line on connected planes representations (Centofanti, 2008).

Here, the traditional representation in 2D orthogonal projection is used in association with GIS mapping and archiving capabilities. The limits consist in the fact that the GIS and interfaces of the different planes are separated and the data are obviously not georeferenced.

3.1.3 A webGIS for the management and the dissemination of CH data

A Virtual Research Environment (VRE) dedicated to the exploitation of intra-site CH data was developed in previous research (Meyer et al., 2007). The information system produced is based on open-source software modules on the internet. The system provides the opportunity to perform exploratory analyses of the data, especially at the spatial and temporal levels (Figure 45 - Figure 46). It is compliant to every kind of cultural heritage site and allows the management of diverse types of data. Some experimentation was done on sites managed by the Service of the National Sites and Monuments of Luxembourg.

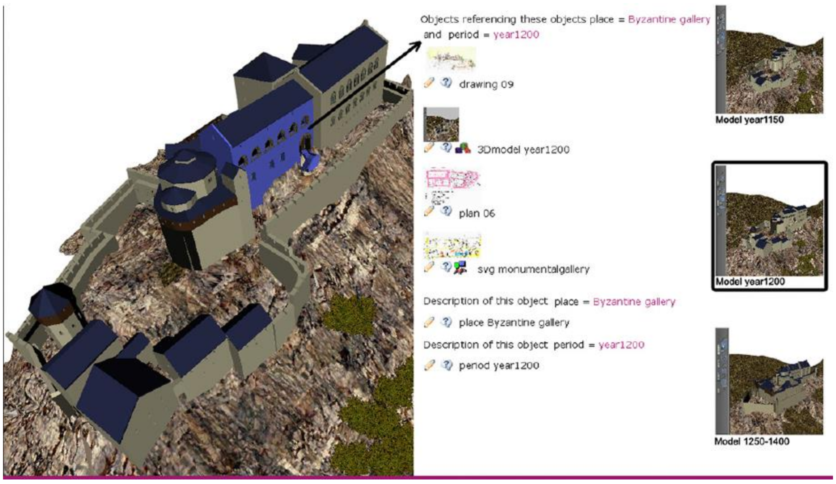


Figure 45. Interactive VRML interface allowing access to data referencing and being referenced by the location clicked. Example of the Vianden castle site (Meyer et al., 2007).

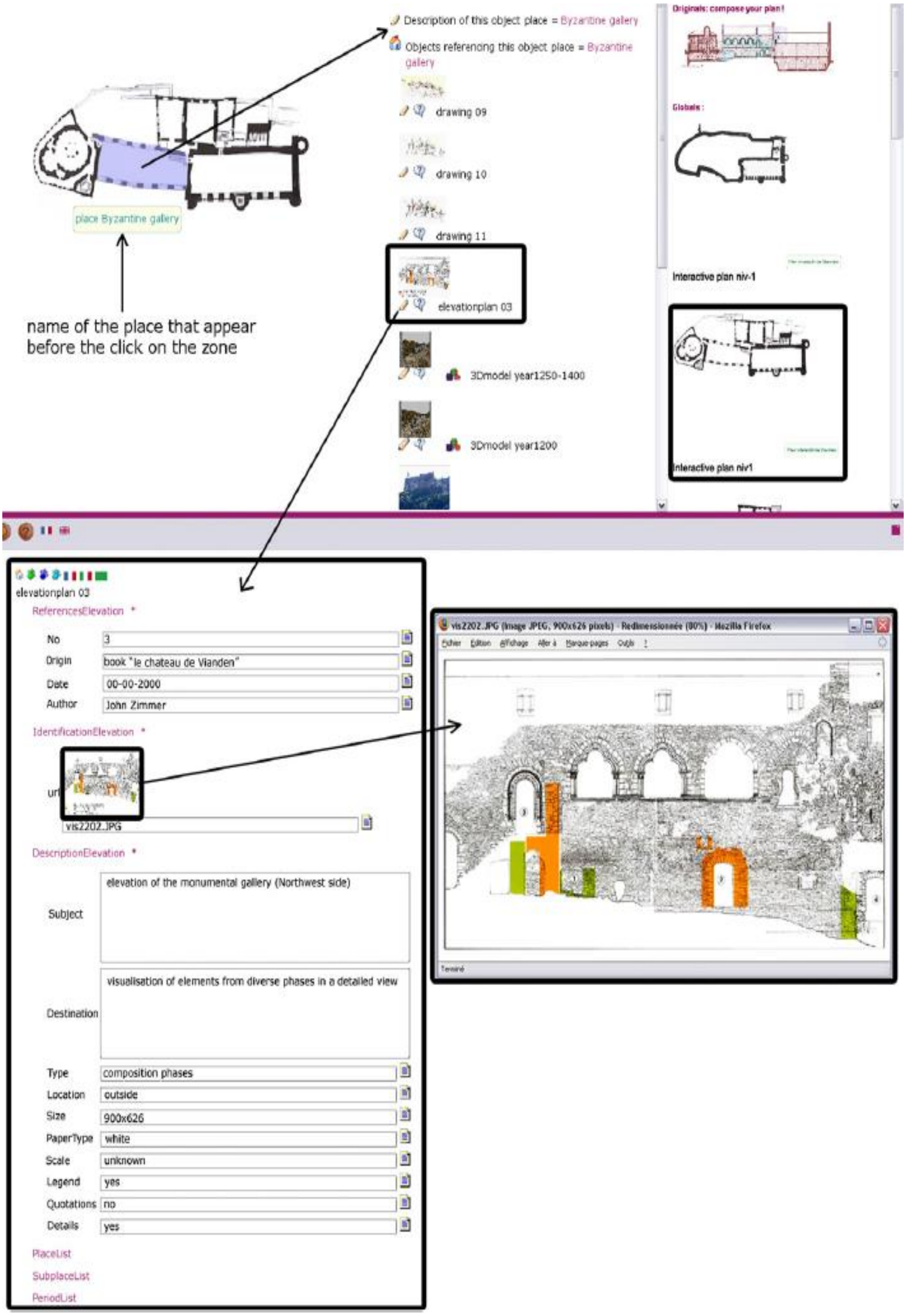


Figure 46. Interactive SVG interface and the correlated documents that can be directly opened afterwards. Example of the Vianden castle site (Meyer et al., 2007).

In addition to the data management possibilities explained before, the user of the VRE can generate his own models and composition plans due to multi-criteria data entry forms written in PHP. He has the choice to select one or several places, and one or several historical phases for which he wants to generate 'on-the-fly' 3D models or 2D drawings (Figure 47), to see the evolutions of the site in time.

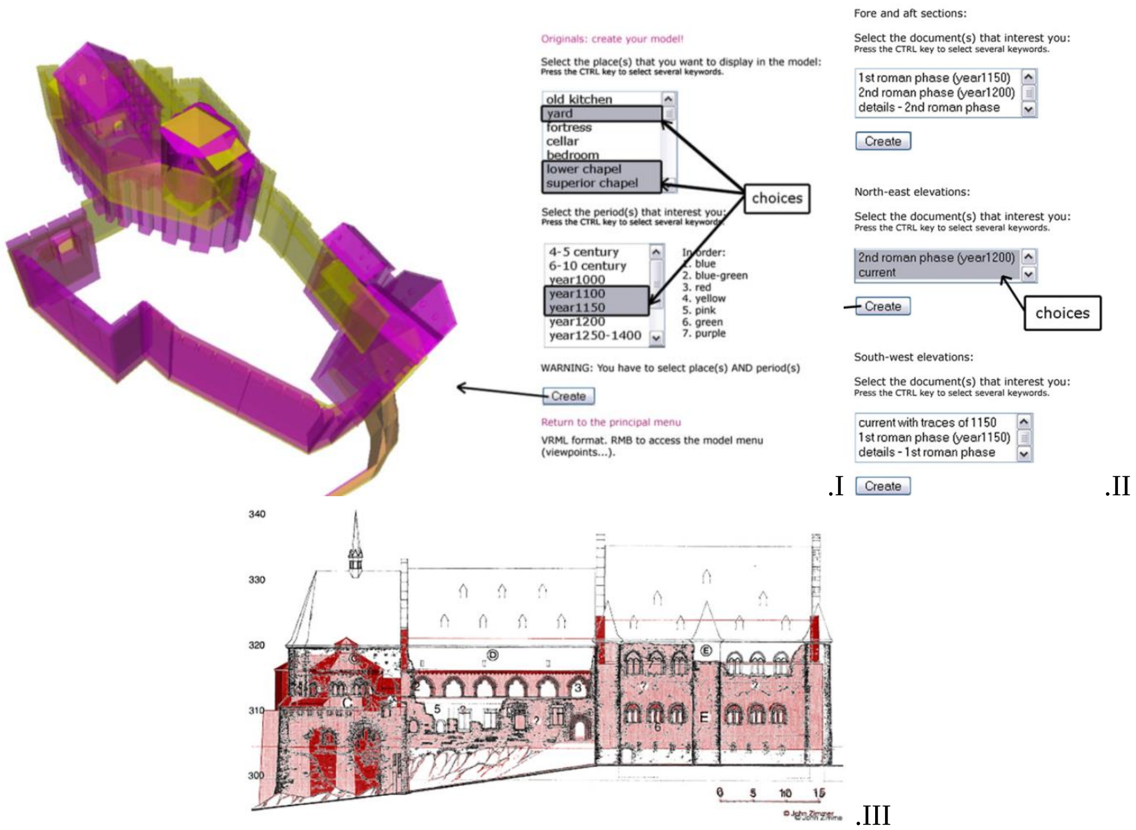


Figure 47. (I) Process to generate 3D models 'on the fly'; (II) process to generate composition plans; (III) 'on the fly' visualization. Example of the Vianden castle site (Meyer et al., 2007).

Also in this case, a specific application was programmed. The reference to standards is not cited.

3.1.4 Image data archive

(Cerutti et al., 2015) proposes the use of GIS systems in order to store architectural images useful for preservation aims. A way for managing close range image data with the use of the open source software QGIS connected to PostgreSQL is described. This tool allows the recording and management of many types of data, including images. They are stored and managed in PostgerSQL and are visualized by means of the database connection to QGIS, where query results and georeferenced image visualizations are displayed: the resulting project contains the image and its attributes about the conservation state and the heritage management activities.

The georeferenced repository is structured in archives able to manage heterogeneous data: the spatial archive, the images archive, and the one regarding CH conservation state. These are queried singularly or by means of an integrated process, according to the aims of the analysis, and they offer, through the images and attributes, a highly detailed reading of geometric and thematic information in the GIS environment (Figure 48), useful also to non-expert users of the system. The database visualization is also realized through loading the PostgreSQL database on a web page by starting from a PHP script connection (Figure 49).

The repository of georeferenced images collects different targeted images, including the ones acquired to generate dense 3D models with low-cost techniques, and it is possible to query the images of the object or images of the points model, obtained by an image matching technique.

The results are different information levels organized according to a multi-scale approach, from territorial to architectural scale, with the displaying of degradations starting from images and 3D metric models.

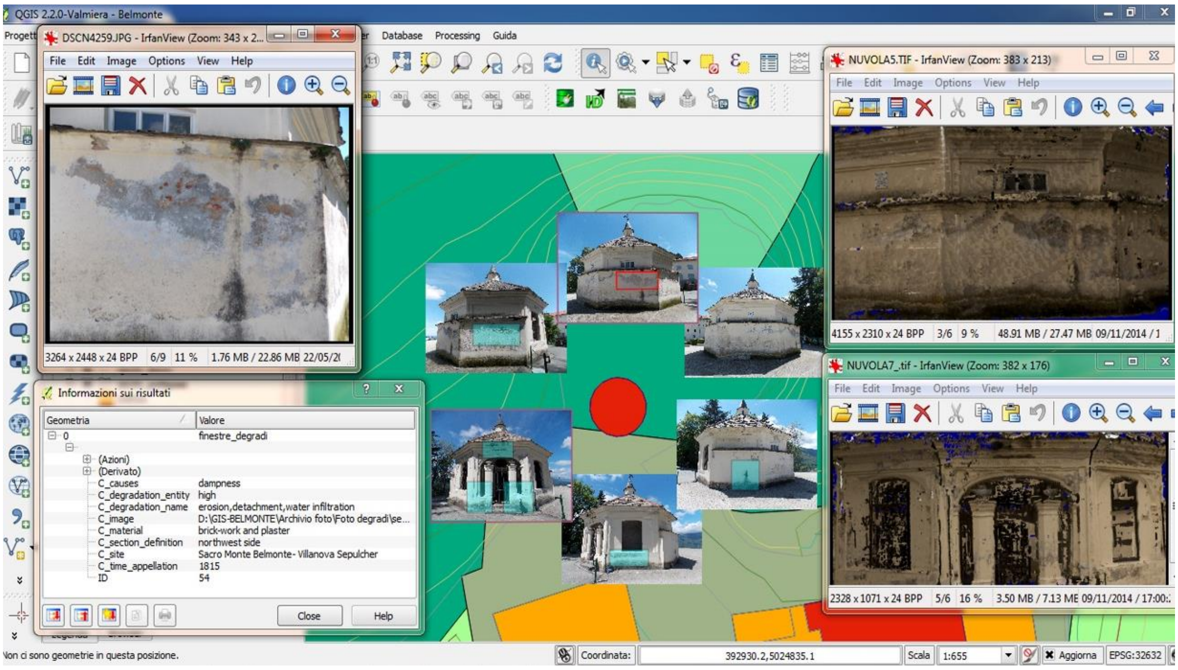


Figure 48. Analysis of materials' degradation from images and 3D metric model in the Belmonte case study (Cerutti et al., 2015).

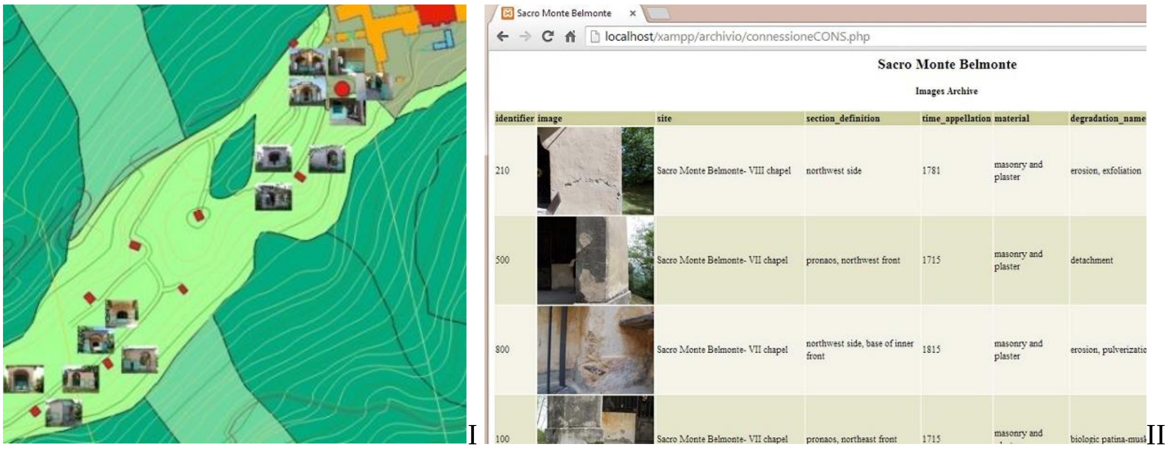


Figure 49. (I) Thumbnail images of Belmonte Sacro Monte (UNESCO Heritage) buildings in GIS environment; (II) PHP Web page connected to PostgreSQL DB (Cerutti et al., 2015).

3.2 GIS as a support for a preservation project

3.2.1 ARKIS project

Its development started in 2000 by the ITABC (Istituto per le tecnologie applicate ai beni culturali) of the CNR (Consiglio Nazionale delle Ricerche) as a project for the documentation of the Roman Theatre in Aosta (Italy)[57]. It is now linked to the Europeana portal, and has an internet domain[58] whose access is not public accessible.

The Architecture Recovery Knowledge Information System (ARKIS) and its evolution ARKIS – NET, is an information system to serve as an auxiliary tool for the organisation, representation and utilisation of knowledge of data in the recovery of historical buildings (including the architectural item, its immediate context and its territorial location) (Salonia, Negri, 2003). It aims to disseminate high-end heterogeneous data (geometric, raster and vector, descriptive, alphanumeric organized in RDBMS and as text, images, in various formats), organised and represented in GIS form, and mapping services via the internet. Clients may display, query, and analysis information in an easy-to-use web browser, and also integrate local data sources with internet data sources. The system is structured as a modular tool consisting of several sub-systems (different levels of knowledge through different degrees of detail), aimed at guaranteeing the organization, representation, and utilization of the knowledge itself. The metric basis support of the mapping is the canonical forms of graphical 2D representation (plan views, elevations and sections), as a geographic area included inside a map.

The GIS functions also regulate all the selection and query operations, which are based on geometric or topological criteria, as well as on logic-arithmetic expressions. Overlay operations (topological cross-referencing) allow the intersection between the different databases to be achieved by superimposing the various graphic-rendered topics and the relative tables of attributes, at the same time ensuring functionality of analysis among the different coverages (information layers) by combining together elements having different common attributes (Figure 50).

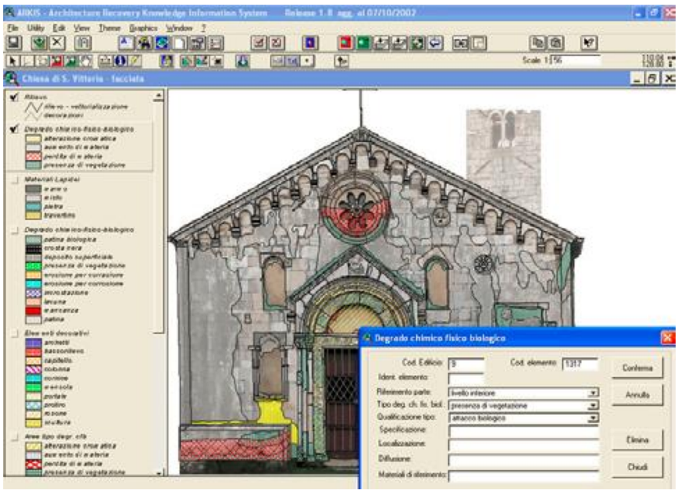


Figure 50. ARKIS: consultation operation of the archive including the decay situation of a study object (Salonia, Negri, 2003).

[57] http://www.cnr.it/istituti/FocusByN.html?cds=098&nfocus=3
[58] www.arkis.it

The system is user-friendly and can be used as a support and guiding operating tool by various different bodies, in particular the superintendencies. However, it is not sufficient just to be an operating tool, considering the huge potential possessed by the system. It can be shared by several users in a network for planning, comparison, discussion and knowledge.

It is limited to 2D representation; therefore, often a local system is used, in order to manage façade plans.

3.2.2 SICaR Information System

SICaR w/b[59] (Sistema Informativo per la documentazione geo-referenziata in rete dei Cantieri di Restauro) was developed within the OPTOCANTIERI project (2003-2004) as a tool to plan and manage restoration work, mapping and geo-referencing data over an exact and measurable model of the artwork (Baracchini et al., 2007). The project was financed by Regione Toscana (Italy) for involving small and medium enterprises in the documentation and preservation process of cultural heritage. The Pisa Superintendence BAPSAE was a partner. It was then experimented by the Ministry during the project ARTPAST and a subsequent one, 'REARTE' (Restauri in Rete). It is the evolution of the Akira GIS server, which was developed with the consultation of the Scuola Normale Superiore of Pisa and the control of the Istituto Centrale del Restauro (ICR), for the project of the stone restauration of the Pisa Tower.

SICaR is an internet-based system that can be used both to consult and insert data. It can interoperate with any data and information available on the web (ICCD catalogues, Risk Map of cultural heritage, pre-existent documentary archives, etc.); it allows the mapping of areas and connects them to forms describing scientific analyses, typologies of deterioration and related repairs activities (Figure 51). It estimates surfaces, costs, and more statistics related to the data.

The managed heterogeneous data are structured in a relational database. They include methods, materials, state of conservation, instruments, chemical and physical analysis, stratigraphy, and so on, to be cross-referenced with further data, such as text, images, videos, etc. All the spatial data are measurable also from the platform and are georeferenced. Open source solutions, a standardized lexicon and a unified structure of restoration data are used for fostering standardization, interoperability and reuse of data and spreading of the system[60]. The data can be directly inserted in the system by the operators on the working site, being the platform available online and enabling editing.

The system[61] is available also as a knowledge base, since operators can query the system for comparing solutions in similar cases and on similar artworks, for serving as support in the preservation project process[62].

In 2011 it was indicated by the Italian Ministry of Cultural Heritage and Activities as the tool to be used for the management of restoration works (Circolare 31 della Direzione Generale per il Paesaggio, le Belle Arti, l'Architettura e l'Arte contemporanee, 2011) (MiBAC, documentation of a technical seminar on SiCAR, Parma 10-11 April 2013).

59 www.liberologico.com/sicar
60 http://www.artpast.iccd.beniculturali.it/fileadmin/ template/allegati/rearte/SICaR_scheda_tecnica.pdf
61 http://sicar.beniculturali.it:8080/
62 http://www.sbappsae-pi.beniculturali.it/index.php?it/300/sistema-informativo-sicar

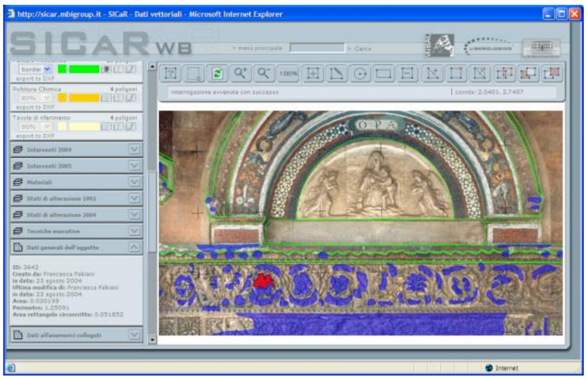


Figure 51. SICaR w/b interface: mapping of interventions on the San Ranieri portal in the Pisa Cathedral (Baracchini et al., 2007).

3.2.3 A standard-based modelling of a multi-view GIS for preservation project support

For this research (realized towards the solution for standardisation finally achieved in this thesis) a GIS structure is modelled by integrating different spatial and thematic standard data-models in order to effectively represent the information as a support for a preservation project (Cerutti et al., 2015). The goal was to provide a system in which information concerning the architectural elements and their measurable morphology, material decay, repair interventions and non-routine maintenance works can be related with the purpose of better coordinating the planning phases of restoration and monitoring activities. Interoperability was a primary issue, therefore available standard data models and open source software (OSS) were used.

The main purposes of the work include the mapping of materials and pathologies of the fronts (one is taken as example) (Figure 52) and the analysis of the building exploiting the representation from different points of view and levels of detail, as well as considering the relation with the landscape context (Figure 53). These aims have been achieved through the use of different linked maps, so that digital regional maps and orthophotos representing the building fronts, or vector graphic drawings representing features of facades, can be visualized separately while remaining stored in a unique geodatabase.

The conceptual model that structures the GIS (Figure 54) is composed by choosing the entities from some useful standards concerning the application field (the standard data model for cartography CityGML, the ontologies for cultural heritage issues CIDOC-CRM, MONDIS). They all follow an object-oriented structure, which offers several advantages in representation (such as the possibility of managing inheritance). This is the reason why the OSS object-relational DBMS PostgreSQL with its spatial extension (PostGIS) was chosen. For managing, analysing and visualising the spatial data, some graphical interfaces were used in QGIS, which is an OSS with common functions of a GIS software. More than one QGIS project was interfaced with the unique database in PostgreSQL-PostGIS, permitting the storage of all the data in a unique archive without duplication. Moreover, in the PostGIS database more than one attribute having a type “geometry” can be inserted for each entity, so that it is possible to store one geometry per level of detail (they can be imported as separate layers in QGIS) for having a multi-scale representation without data duplication.

On the other hand, the data can be analysed as a whole in the PostgreSQL DB, which contains all the records. The different interfaces are linked to each other for permitting the passage from one to the other reading level. The operations on the data, the computing of statistics and the possibility to execute meaningful spatial analysis on the archived data can be of interest for optimizing the phases and some analysis of the preservation project.

Figure 52. (I) GIS representation of the 'Materials' values mapped on the façade and related table in PgAdminIII. When possible, taxonomy values are used, preferably kept from the NORMAL documents; otherwise, other affirmed bibliographic sources were used (Carbonara, 2004). (II) GIS representation of the "ManifestationOfDamage" values mapped on the façade and related tables (Cerutti et al., 2015).

Figure 53. (I) GIS layout representing the data of the studied object (that is, the ex-convent Building annexed to the 'Chiesa del Colletto') in its context. From this regional scale map it is possible to link another GIS project where another view at a major scale is represented (the

projection is the red line). (II) QGIS visualisation and the values of attributes of the directed query line (projection of the analysed front). Through 'Actions' it is possible to open the other project (Cerutti et al., 2015).

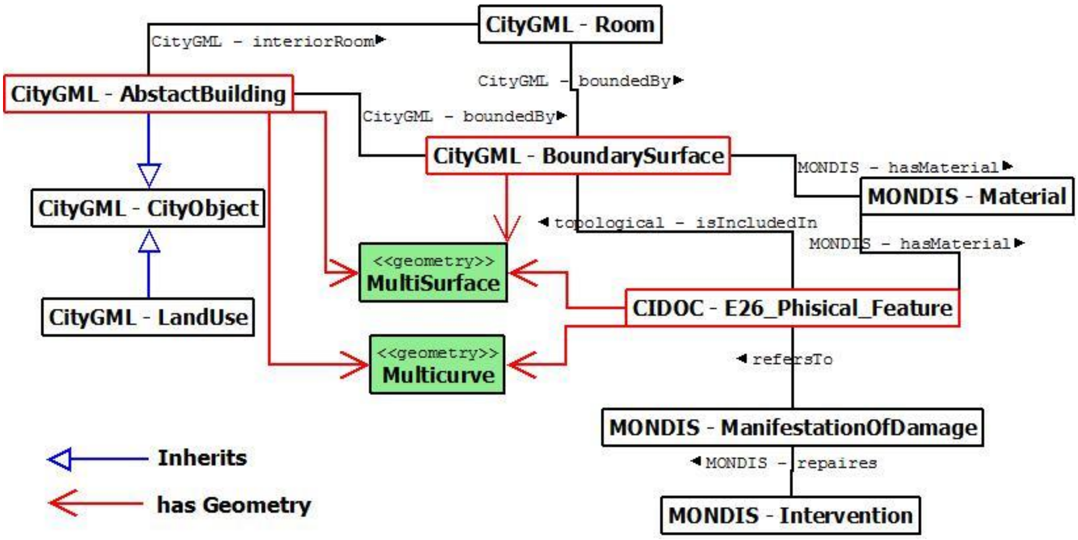


Figure 54. Conceptual model derived from the integration of existing standard data models used for the application (Cerutti et al., 2015).

3.2.4 NUBES web-based documentation platform for mapping and displaying cultural heritage surfaces

The goal of this research (Stefani et al., 2014) is to supply authorities of an architecture with a tool for scientific monitoring and decision support, based on the acquisition of alteration mapping and on the estimation of degradation patterns, in order to permit rational programming operations of restorations. An approach for digital documentation about the conservation state of buildings is proposed. It is based on the connection between a structured 3D model and structured 2D mapping data. In the NUBES web platform (open source web platform for describing, analysing, understanding, documenting, and sharing heterogeneous data at the architectural scale), a specific interface has been implemented, permitting the displaying and cross-referencing of 2D mapping data on the 3D model in real time, by means of structured 2D layer-like annotations concerning stone degradation, dating, and material (Figure 55).

The process of enrichment of information is based on three main steps:

- the uploading of textures belonging to digital models;
- the recording of stone alterations and their display on the 3D model (a free web-based, JavaScript-driven SVG editor permits the drawing/modifying of alterations in specific layers; alterations are saved directly in the database and uploaded on the fly in the 3D viewer);
- the display on the fly of the crossing map in the 3D scene (the 3D scene constitutes a privileged environment in which to consult cartographies, as architectural discontinuities and points of joining between the different elements are more readable).

The platform also allows the insertion of specific points in the 3D scene such as those corresponding to the photographs of details, iconography, meteorological data, lapidary signs or sensor data.

Patterns can be drawn on specific hierarchical layers concerning three main categories of analysis on stones: degradations, dating, and material. Each category is structured in hierarchical sub-layers, each of them containing sub-types. As an example, the stone degradation category is divided in biological, physico-chemical and mechanical sub-types.

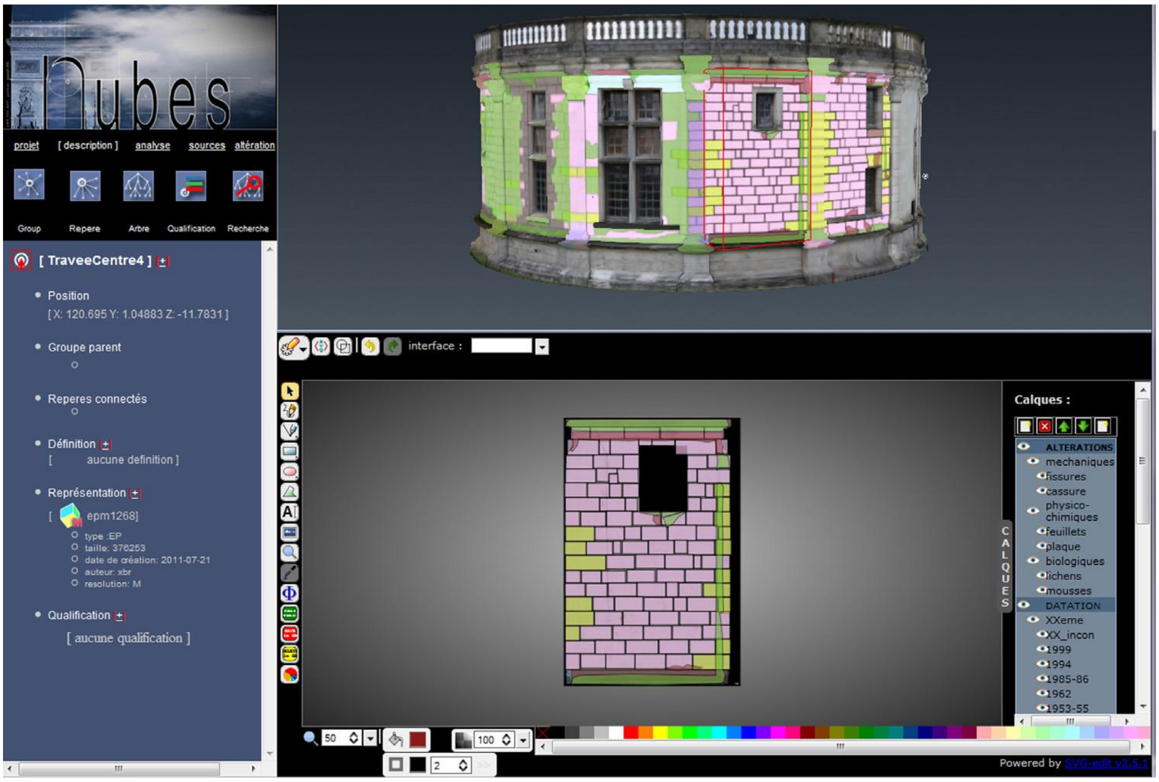


Figure 55. Interface of the NUBES platform specific to the stone degradation pattern documentation (Stefani et al., 2014).

An interesting research was also performed for obtaining an analogous platform by means of open source and free software (Guarnieri et al., 2010).

3.3 Stratigraphic Analysis in GIS

3.3.1 A 3D digital workflow for archaeological intra-site research using GIS

(Katsianis et al., 2008) proposes a formal data model and complete digital workflow for the documentation of the archaeological stratigraphy analysis process in 3D. The subject here is not properly the architecture, but, since 3D approach are used, it could be of interest nevertheless. The entire digital process has the advantage of being implemented on a single software platform. In addition, the combination of formal ontology and custom object-oriented programming enables a suite of techniques for exploratory data analysis and stratigraphic interpretation.

The focus in on the combination of three critical elements:

- the development of an explicit data model for georeferenced archaeological data;
- effective recording and handling of spatial entities in 3D (Figure 56);

- the development of 3D tools for intra-site analysis (Figure 57).

The described approach is based on an explicit semantic and geospatial model and a relatively transparent development environment (ESRI ArcObjects[63]) that allows communication with different software platforms and is potentially open to further adaptation and customization by other users (it is, however, proprietary software).

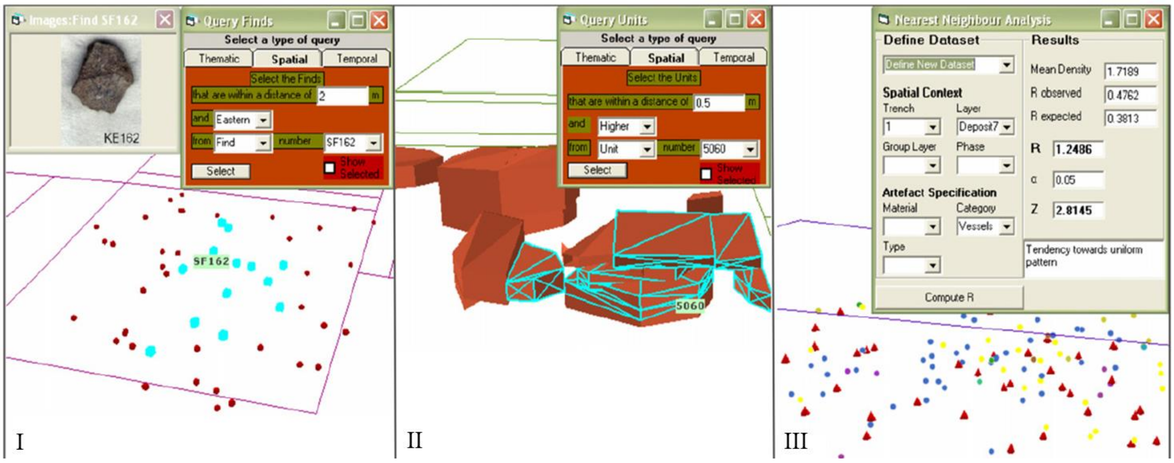


Figure 56. Data exploration in 3D space: (I) selection of finds by 3D distance, (II) selection of excavation units by 3D distance, (III) applied nearest neighbour analysis (Katsianis et al., 2008).

[63] ArcObjects is a set of platform-independent software components, written in C++, that provides services to support geographic information system (GIS) applications on the desktop in the form of thick and thin clients and on the server. The language chosen to develop ArcObjects was C++; in addition to this language, ArcObjects makes use of the Microsoft Component Object Model (COM). COM is often thought of as simply specifying how objects are implemented and built in memory and how these objects communicate with one another. While this is true, COM also provides a solid infrastructure at the operating system level to support any system built using COM. On Microsoft Windows operating systems, the COM infrastructure is built directly into the operating system. For operating systems other than Microsoft Windows, this infrastructure must be provided for the ArcObjects system to function. (http://edndoc.esri.com/arcobjects/9.2/net/5bd93a2b-1c00-4927-ab26-5fbe3891a448.htm)

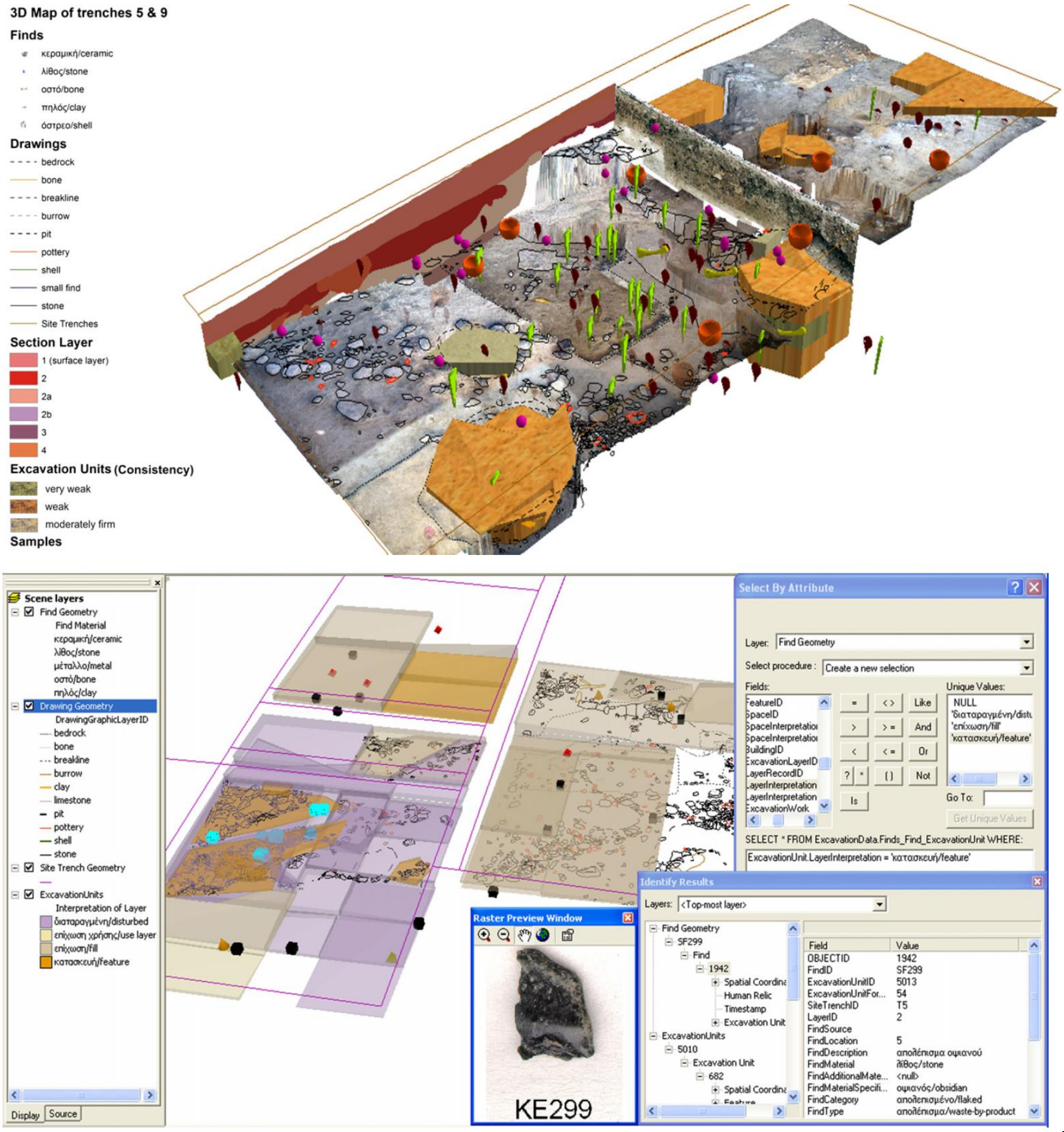


Figure 57. *(I) Combination of trenches, excavation units, finds, drawings, and photos with advanced symbology; (II) thematic data classification, thematic selection and identification of features (Katsianis et al., 2008).*

3.3.2 A GIS for Medieval Archaeology Based on Photogrammetry and Archaeological Database: The Shawbak Castle Project

(Drap et al., 2012) presents an interdisciplinary project (Figure 58) towards a 3D Geographical Information System (GIS) dedicated to cultural heritage with a specific focus application of the Castle of Shawbak (rural medieval settlements in the Middle East). Focusing mainly on stratigraphical analysis of upstanding structures, a two-step pipeline was developed. First, a survey process based on photogrammetry is applied, and then a tool for statistical analysis is utilized. Two main applications are presented: stratigraphy analysis with a Harris matrix (a strategy for classifying a stratigraphic unit in relation to adjacent ones) computed on the fly from the 3D viewer, and statistical tools, clustering operation on ashlars in order to show new relationships between the measured artefacts. All the developments are written in Java by means of the jReality application programming interface which is an Open source Java class library for interactive visualization of a large spectrum of 3D data (Figure 59).

For achieving these goals, two approaches were merged coming from two different scientific fields. From one side an archaeological analysis with a database processing was used: the archaeologists' team have long employed a database designed for this application that is populated by all the archaeological observations and use concepts also utilized by the survey process, such as CF (*Corpo di Fabbrica*, i.e. Building Unit) or USM (standing for *Unità Stratigrafica Muraria*, i.e. Wall Stratigraphic Unit). The second approach is photogrammetry and computer science in order to produce 2D or 3D models. A bridge between the two approaches was developed in order to add semantic data on the cloud of 3D points.

Once the ashlar blocs are measured and linked to the archaeological database (at least by their own USM) there are two important links between them: temporal relations, given by the Harris relationships in the database and structural relations given by the dimensions and position of each bloc or USM (Figure 59).

A 3D viewer was developed to be able to visualize all data used, photogrammetric models, with photographs and the 2D/3D points, ashlar blocs, dense cloud of 3D points, USM and clustering analysis and finally the USM graph used in the Harris matrix. It allows interactivity between the stored data (the data persistence is done with both XML formalism and relational database) and the user (Drap et al., 2009).

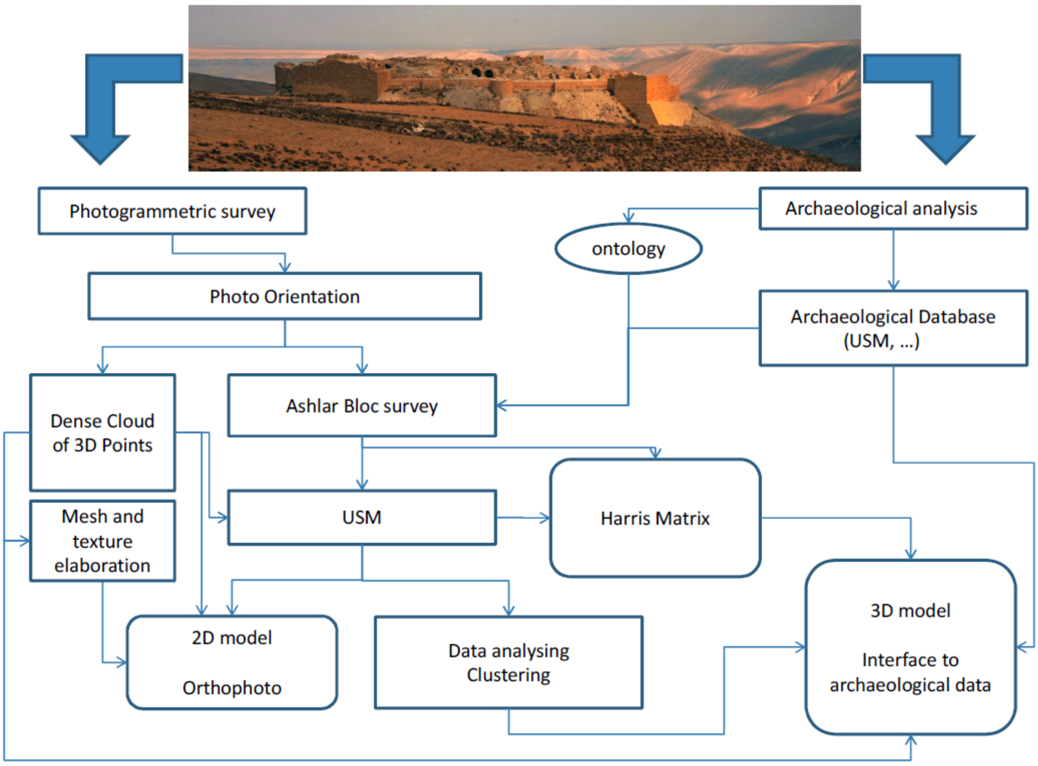
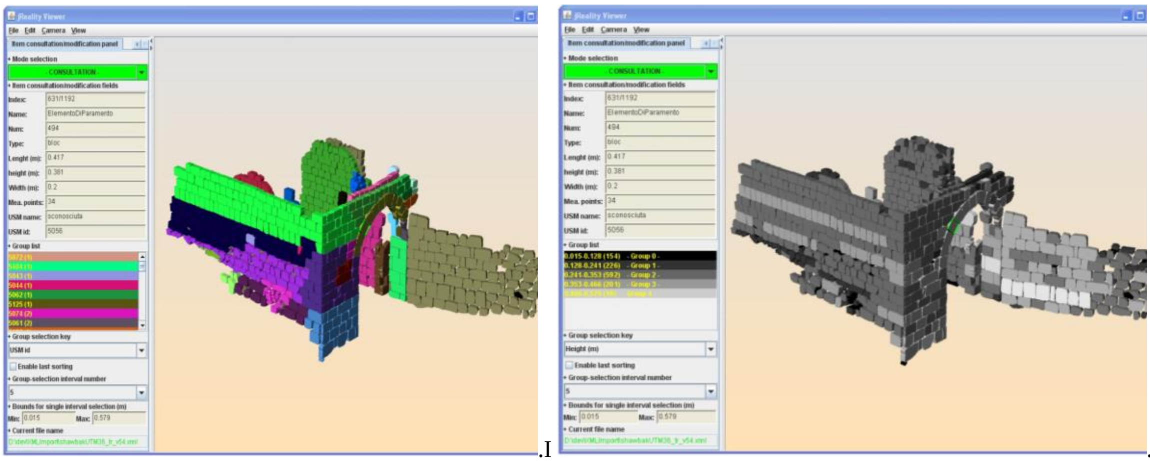


Figure 58. Synoptic schema of the interdisciplinary project, with detail of archaeology and computer science cooperation and a view of Shawbak Castle (Drap et al., 2012).



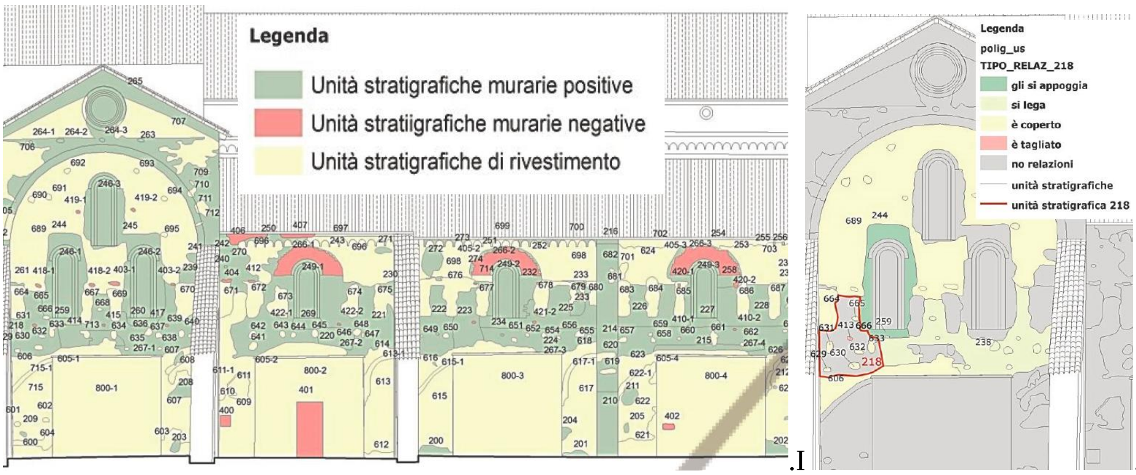
.I .II

Figure 59. I. 3D visualisation with a colour panel chosen according to the block height classification. II. 3D visualisation with a colour panel chosen according to USM (Wall Stratigraphic Unit) (Drap et al., 2012).

Here the 3D component and the analysis capabilities are very effective. Again, a specific platform is programmed, and standards for representation are minimally used.

3.3.3 Data Collection and Management for Stratigraphic Analysis of Upstanding Structures

Stratigraphic analysis, used for study of archaeological excavation, has been adapted and applied to upstanding structures with the aim to reconstruct a building's history (Donadio, Spanò, 2015). The stratigraphic units were mapped on the fronts with the aid of orthophotos and textured 3D models of the façades. The next step consisted in the identification of the temporal relationship between each unit and its adjacent ones. In order to archive and link this kind of descriptive information, an attribute database was built. Through the management of this integrated spatial and thematic information, useful thematic maps and queries (Figure 60) can be performed for the study of the building history.



.I .II

Figure 60. (I) Example of a data thematisation: green shows positive USM (Wall Stratigraphic Unit), red indicates negative USM, yellow is covering units. (II) Result of a select by attribute query about the 218 USM. The system allows highlighting all the USM that have a relationship with it and to thematise features based on the type of relationships they have (Donadio, Spanò, 2015).

3.4 Final comments

The described examples offer many interesting issues: 2D/3D data management, the representation of time, the management of both relational and object-oriented structures, the possibility to include in the representation multiple format resources and to visualize the object from multiple views, the sharing of the information on the web, and the possibility of participation in completing the archive by different operators through the web. All of these aspects are critical for similar representations.

However, some additional points should be integrated in the representation for the aims of this research. First of all, many systems, especially regarding the most detailed representations, are in a local reference system, while it could be important to refer the data to their position in a global system, for permitting better inferences considering spatial issues, as well. Then, the data are often structured following sectorial data models and rarely referencing to a shared standard. Also the link to similar databases is often not present.

The integrations, which are aims of this thesis, concern the use of standard ontologies and data models, which could interface with cartography archives for the enhanced possibility of data retrieval, analysis and reasoning processes. For the same aims, a relation to external correlated databases and vocabularies is pursued.

A final remark regards the implementation point of view: the more charming and effective results for the archives consultation interfaces and the possibility to include advanced functions in the systems (queries, rules, the passage from 2D to 3D, object-oriented structures and so on) are based on expressly programmed applications, since existing tools still have weaknesses related to these aims. Related to this, some research is being developed in this vein (Chaturvedi, 2014; Chaturvedi et al., 2015; Ionifescu-Enescu, 2007).

4 CH DIGITAL INVENTORIES AND 3D-MODELS WEB ARCHIVES

The institutions managing cultural heritage preservation and documentation have found digital technologies to be very useful tools to pursue their goal.

The tendency to develop digital archives of the world heritage is now widespread, and the standards have an important role in this process, as explained in Section 6.2. Obviously, semantic web technologies (Chapter 7) can also help in effectively archiving the resources, and therefore they are increasingly used.

Still without considering standardisation, but best exploiting the available 3D survey techniques, some 3D-models about cultural heritage are developed and published on the web.

In fact, the 3D modelling tools and survey methods permit the obtaining of dense, accurate, and detailed 3D models in a reduced time and employing few economic resources to simplify the massive digitalization of objects, for being used as a support for communication, analysis, research, education, tourism, and all the issues exposed in previous chapters, including preservation.

For these reasons and with these goals, some organizations dealing with cultural heritage began to build digital archives that can be browsed and explored through the web. For example, the Smithsonian Institution (Section 4.3) archive and the archive developed by the CyArk Project (Section 4.4) are reported in the next chapters.

The browsing and analysis of the digital documents, in such broad archives, could have great advantage from the definition of a shared semantics of the represented objects. The use of ontologies (Chapter 5) for the management should be a solution to this task, and, also, could assist in reaching the goals of interoperability, communication, education and analysis.

First promoters of the employment of new technologies for documentation and archiving aims are some institutions, for example, the Getty Institute and the Smithsonian Institution, which are presented in the next sub-sections.

4.1 The Getty institute

The Getty Institute has been one of the most active private institutions in the documentation of research and the preservation of cultural heritage, since 1953. It is active in promoting research about preservation issues and sharing knowledge about cultural heritage[64].

A part of the Getty Institute’s activity is directed to the formulation of standards for cultural heritage and the realization of digital inventories. They are often compliant with existing standards, for interoperability reasons. Since the main Getty products, which are of interest in this research, exploit linked open data and further technologies explained in Chapter 7, they are described in that chapter.

***Mission**: the J. Paul Getty Trust is a cultural and philanthropic institution dedicated to critical thinking in the presentation, conservation, and interpretation of the world's artistic legacy.*

Through the collective and individual work of its constituent Programs (the Getty Conservation Institute, the Getty Foundation, the J. Paul Getty Museum, and the Getty Research Institute) it pursues its mission in Los Angeles and throughout the world, serving both the general interested

[64] www.getty.edu

public and a wide range of professional communities with the conviction that a greater and more profound sensitivity to and knowledge of the visual arts and their many histories are crucial to the promotion of a vital and civil society.

Vision*: the Getty and its four programs are dedicated to the proposition that works of art are windows onto the world's diverse and intertwined histories, mirrors of humanity's innate imagination and creativity, and inspiration to envision the future.*

To this end, the Getty works to:

- *enhance understanding of art through innovative, digitally driven research, shared for the benefit of the widest possible audience;*
- *strengthen and broaden our collections to provoke the curiosity of scholars and visitors alike;*
- *chart a new course for how art, humanities, and cultural and public policy can together foster a more inclusive, vibrant civil society;*
- *serve as a convener and catalyst in the cultural life of Los Angeles;*
- *offer transformative experiences for visitors to our collections, gardens, and facilities at the Getty Center, the Getty Villa, and the Getty's presences online, free of charge now and forever.*[65]

4.2 The Smithsonian Institution

The Smithsonian Institution, founded in 1846 in the USA, is devoted to the preservation of heritage, and to promoting research, culture and educational activities, by administrating museums and galleries and supporting a number of related activities.

Mission: *the increase and diffusion of knowledge.*

Vision: *shaping the future by preserving our heritage, discovering new knowledge, and sharing our resources with the world.*

Values:

- *discovery (explore and bring to light new knowledge and ideas, and better ways of doing business);*
- *creativity (instill our work with imagination and innovation);*
- *excellence (deliver the highest-quality products and services in all endeavours);*
- *diversity (capitalize on the richness inherent in differences);*
- *integrity (carry out all our work with the greatest responsibility and accountability);*
- *service (be of benefit to the public and our stakeholders).*

Priorities:

- *four grand challenges (focus on the four grand challenges outlined in the Smithsonian Strategic Plan*[66]*);*
- *broadening access (digitizing our collections*[67]*, exploring next-generation technologies, and improving the visitor experience);*

[65] The italic text cites http://www.getty.edu/about/whoweare/mission.html

[66] https://www.si.edu/Content/Pdf/About/SI_Strategic_Plan_2010-2015.pdf

[67] http://dpo.si.edu/

- *revitalizing education (serve as a laboratory to create models and methods of innovative informal education that link to the formal education system);*
- *crossing boundaries (interdisciplinary consortia spark innovative research and educational programs around each of the four grand challenges);*
- *strengthening collections (develop collections plan to support Institution-wide initiatives);*
- *organizational excellence (strengthen organizational services that allow us to deliver on our mission);*
- *measuring performance (establish performance indicators that will specifically and annually measure progress toward our goals).*[68]

4.3 The Smithsonian X 3D

An important digitalization activity is realized in the collections of the Smithsonian Institution, using different technologies, from digital scanning of paper documents to the realization of digital archives of 3D models of the collected objects.

Since a very small fraction of the collections is shown to the public, the realization of an online digital archive is critical to reach new audiences and to be available for early analysis by researchers, educators and curators. Furthermore, digital collections allow different audiences of users to go on their own journeys of learning, possibly revealing further knowledge[69]. The models can be explored (Figure 61) and downloaded from the web site[70], and are capable of being printed by 3D printers for different aims, with education being the priority.

As it is possible to see from Figure 61, the models are very detailed, and could serve as support for quite complete remote analysis (for example, it is possible to see the object characteristics, degradation, materials, etc. in a quite exact way only by seeing the textured model on the web).

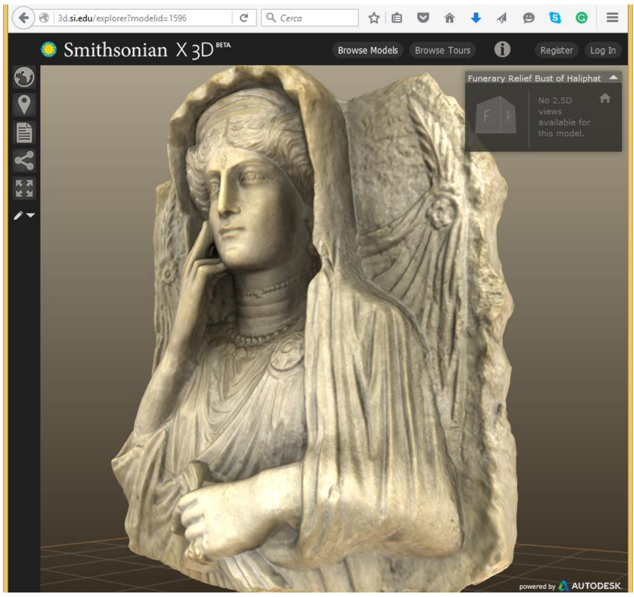


Figure 61. One of the 3D models archived by the Smithsonian X 3D, explored on the web.

[68] The italic text cites http://www.si.edu/About/Mission
[69] http://3d.si.edu/about
[70] http://3d.si.edu/browser

4.4 CyArk

CyArk, was founded in 2003 to ensure heritage sites are available to future generations, while making them uniquely accessible today. CyArk operates internationally as a non-profit organization with the mission of using new technologies to create a free, 3D online library of the world's cultural heritage sites before they are lost to natural disasters, destroyed by human aggression or ravaged by the passage of time.[71]

The CyArk 500 Challenge is CyArk's ambitious goal to digitally preserve 500 cultural heritage sites within the next five years. CyArk and its partners are on a mission to save these cultural heritage sites digitally before more are ravaged by war, terrorism, arson, urban sprawl, climate change, earthquakes, floods, and other threats.[72]

As it is possible to see in Figure 62, besides showing the digital model, the CyArk platform also provide some thematic details about the represented object. These, however, can be increased and improved with a major detail and more complex semantics.

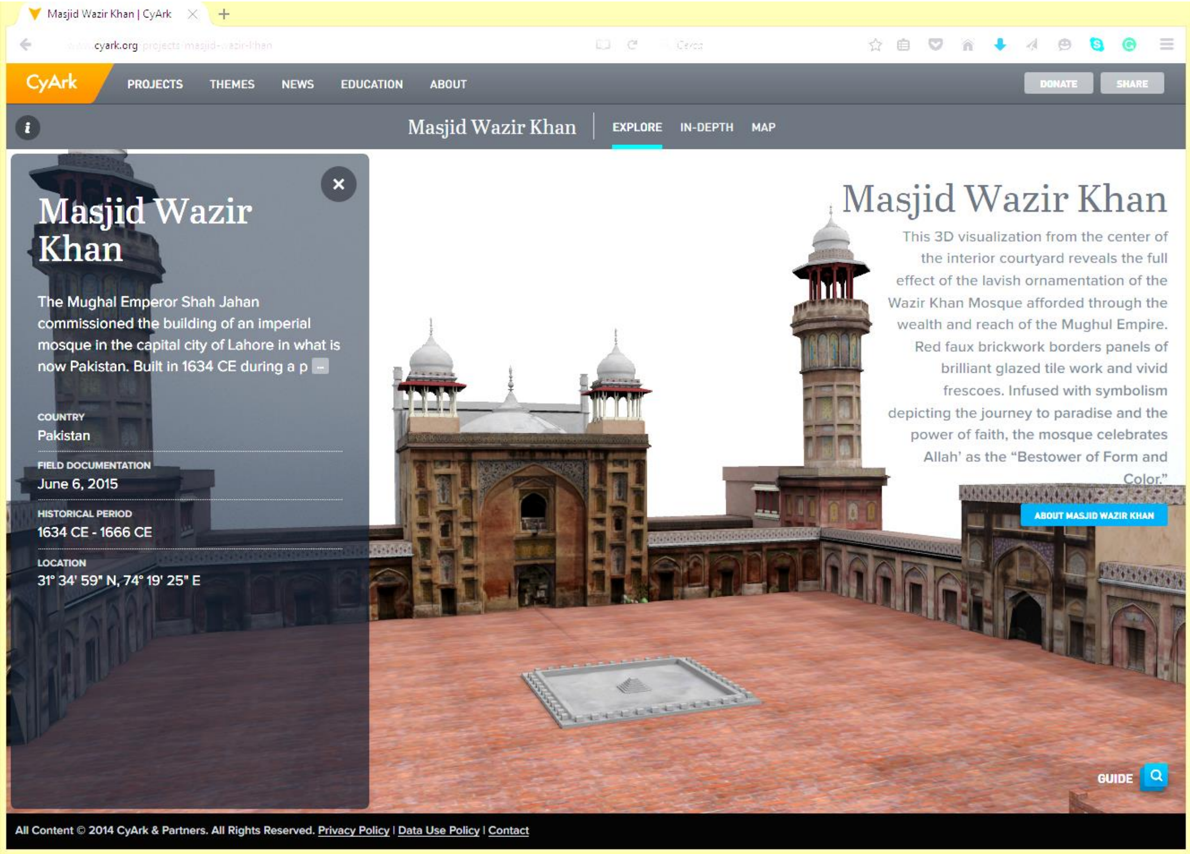
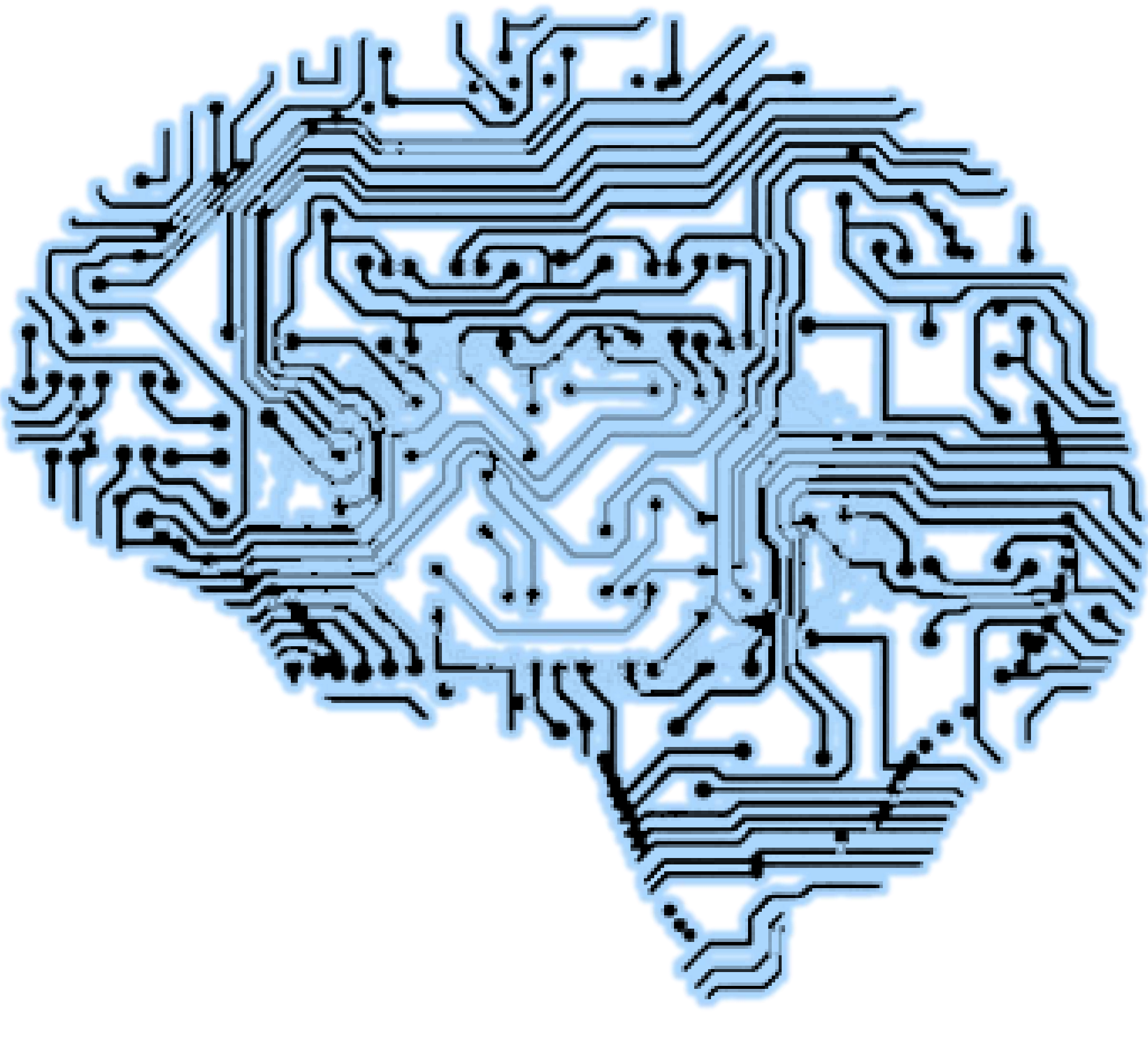


Figure 62. Example of the CyArk digital resource, including the texturized 3D model, a brief descriptiuon about the object and its history, and some data about its location. It can be accessed via the web (http://cyark.org/projects/masjid-wazir-khan).

[71] The italic text cites http://www.cyark.org/about/
[72] The italic text cites http://www.cyark.org/about/the500/

SECTION 2 – THE METHOD



(Adapted from http://www.spettakolo.it/wp-content/uploads/2015/02/viv.png)

5 THE DOMAIN ONTOLOGIES

5.1 The philosophical ontology

The term 'ontology' comes from the Ancient Greek ὄντος, òntos (genitive singular of the present participle ὤν of the verb εἶναι, èinai, 'to be') and from λόγος, lògos ('discourse'), so, literally it means 'discourse on the being', but it could also derive directly from τά ὄντα, that is 'the beings', which is variously interpretable basing on different philosophical positions.

It is the study of 'what there is' (Quine, 1948). Ontology was born with Aristotle, which dealt with the study of the nature and structure of reality and things *per se* in his *Metaphysics* (as was defined the Aristotelian work following the 'physikà' by Andronico di Rodi). Aristotle defined a philosophy as a 'science of the whole of the reality', moving from the sensible world to transcend it for finding the universal fundamentals (Masolo et al., 2003). The heart of the work is the study of the 'being qua being'. It is the science of kinds and structures of objects, properties, events, processes and relations in every area of reality (Smith, 2003).

To this Aristotelian original conception of ontologies, another one can be added, which consists in the study of conceptualizations intended as abstractions of the reality and its categorisation as filtered by our (or our cultural) mind activity and structure (Masolo et al., 2003). Immanuel Kant was the promoter of such a conception in his *Critique of Pure Reason* (Kritik der reinen Vernunft, 1781), and changed in following the development of his philosophy for considering culture and language as determining factors in forming schemas in the human mind (to be added to innate categories of human thinking) (Masolo, et al., 2003).

In a global semantic web perspective, this distinction has to be kept in mind for permitting a conceptualization that can be understood worldwide. More exhaustive analysis of ontologies and their application can be found in a number of works (e.g. Masolo et al., 2003, Mulligan, 2013). Moreover, relations can be referred to a somewhat different part of philosophy, and specific papers analyse them (Mulligan, 1998).

The interpretation of Aristotelian ontological theories for computer science is translated to the study of 'content qua content' and the consequence that this implies are described by a number of studies (Smith, Ceusters, 2010).

In the philosophical sense, we may refer to an ontology as a particular system of categories accounting for a certain vision of the world. As such, this system does not depend on a particular language: Aristotle's ontology is always the same, independently of the language used to describe it (Guarino, 1998)

5.2 Applied ontologies

5.2.1 A premise: the logic

Logic and language are essential concepts (often interrelated) for understanding the meaning of the term ‘formal’.

A logic encodes with a precise set of deterministic rules (inference rules) some basic reasoning steps, considered correct by all. A correct reasoning allows to show that a certain knowledge is a logical consequence of a given set of facts; in logically correct reasoning, chains are constructed by concatenating applications of simple inference rules to reach a logical conclusion (Sowa, 2010).

5.2.1.1 Language logic and logical languages

The language is essential to define and communicate the description of the real world (Figure 64). It can be informal (natural language, graphical language, icons…) or formal (logical language, mathematical language, programming language…). Also some mixed languages exist (for example UML, which is important in database conceptual modelling). Figure 65 shows a representation of the languages with increasing formalization levels. In any case, the language theories are based on the following concepts of intension and extension of the meaning.

Intension (concept) is the part of meaning corresponding to general principles, rules to be used to determine reference (typically, abstractions from experience). Intensional knowledge is usually thought to be ‘timeless’ and unchangeable (Nardi, Brachman, 2003). This corresponds to the higher part of the ‘triangle of meaning’ (the concept) (Figure 63).

Extension (object) is the part of meaning corresponding to the effective reference. Extensional knowledge is usually thought to be contingent (Nardi, Brachman, 2003). This corresponds to the lower part of the ‘triangle of meaning’ (object or symbol level) (Figure 63).

By means of the concept associated to the sign ‘cat’, we can correctly interpret this sign in various situations. The sign’s referent is the result of this interpretation.

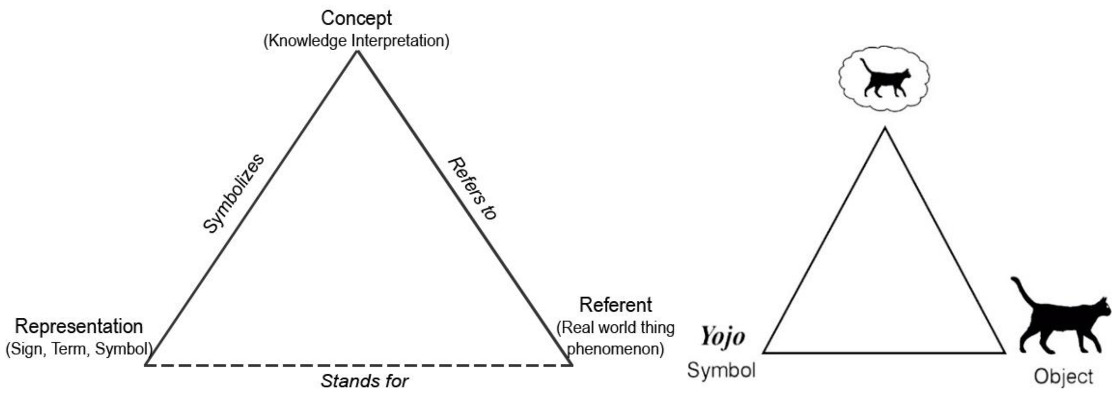


Figure 63. Triangle of meaning[73] (Ogden, Richards, 1993) and an example[74] (yojo is the name of a particular cat).

[73] http://inmyownterms.com/wp-content/uploads/2014/06/semiotic-triangle.png

[74] http://www.jfsowa.com/figs/yojo.gif

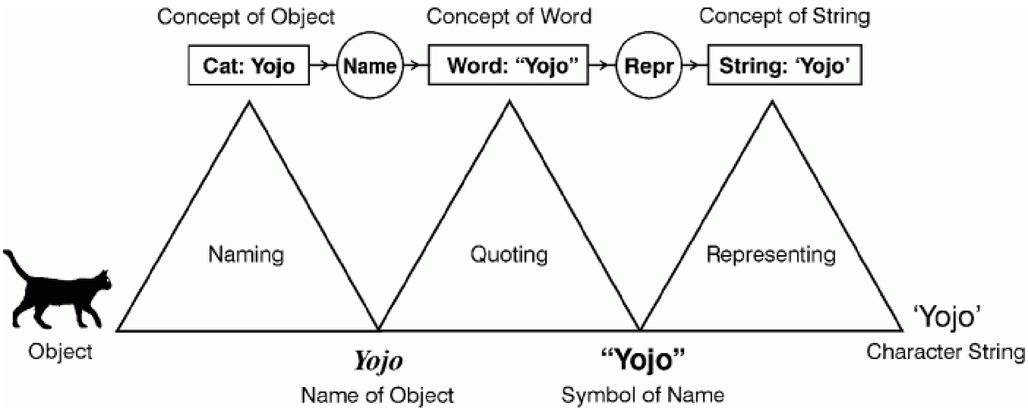


Figure 64. The passages to the encoding of a language[75]

Different interpretations are possible:

- intuitive interpretation (informal semantics): it is the association to any element of the language and interpretation of the real world;
- formal interpretation (formal semantics): a function allows the transformation of the element of the language into elements of the mathematical structure. It is the formalization of the intuitive interpretation (intuitive semantics);
- abstraction: it is the link between the real world and a mathematical structure.

It is essential, in order to correctly communicate, to impose constrains to the possible interpretations. The languages for expressing conceptualizations are therefore classified as follows by (Guizzardi, 2010).

From the logical level to the ontological level (Table 1) the following statements can be written (where ∃ means 'exist' and ∧ means 'intersection'):

- logical level (no structure, no constrained meaning)

∃x (Apple(x) ∧ Red(x))[76] MEANS→ a red apple exists

Logical level languages are "flat": they put all predictive terms (e.g. Apple and Red) on the same footing (Guizzardi, 2010).

- Epistemological level (structure, no constrained meaning)

∃x:apple Red(x)[77] (many-sorted logics) MEANS→ a is an Apple with Colour=red (description logics)

∃x:red Apple(x)[78] MEANS→ a is a Red with Shape=apple IT HAS NO SENSE MEANING

Epistemological level languages (Unified Modelling Language - UML, Entity – Relation schemas - ER, Ontology Web Language - OWL) provide ways for elaborating structures that differentiate the

[75] http://www.jfsowa.com/figs/yojo3.gif
[76] A x exists, which is the intersection between an apple and the colour red, that is, it is both "apple" and "red".
[77] A x exists, which is an "apple", and "red" is a characteristic of x.
[78] A x exists, which is a "red", and "apple" is a characteristic of x.

meaning of terms, but do not give a precise criterion to explain why one structure is better than another. This should be motivated by recognized ontological distinctions (Guizzardi, 2010).

- Ontological level (structure, constrained meaning)

Some structuring choices are excluded because of ontological constraints: Apple carries an identity condition, Red does not.

Ontological level languages (UML 2.0) permit the postulation of ontological distinctions and a rich axiomatization prescribing how distinctions can be related. They allow explicitness regarding their ontological commitments (to minimize the 'false agreement problem', and support the user in justifying his modelling choices).

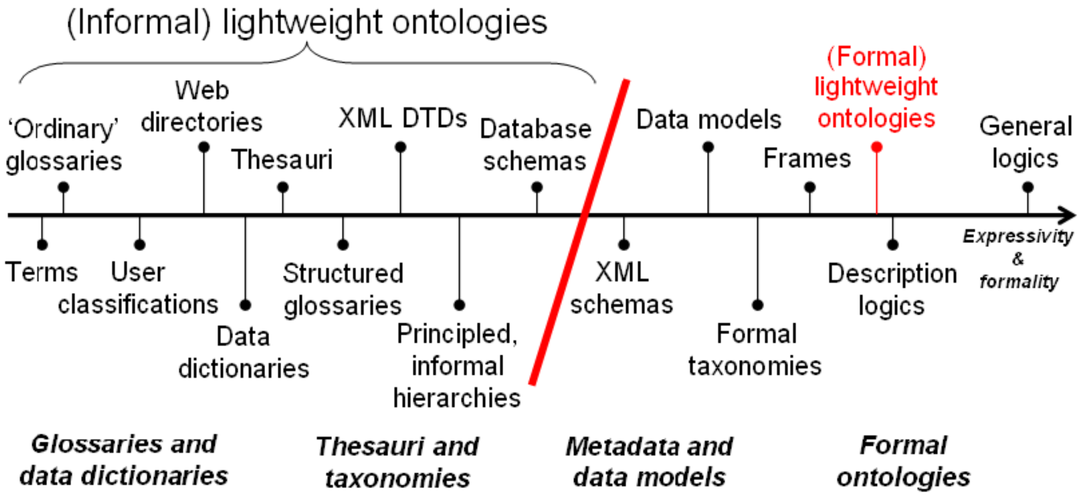


Figure 65. Different approaches to the languages. The ambiguity of the interpretation is reduced in the sense of the arrow. In the left parts, one can speak about 'classifications', which focus on access, based on pre-determined criteria (encoded by syntactic keys), while in the right parts, the meaning of terms and the nature and structure of a domain is the focus[79].

Table 1 Synthesis of the possible levels of a representation (Guarino, 1994).

Level	*Primitives*	*Interpretation*	*Main feature*
Logical	Predicates, functions	Arbitrary	Formalization
Epistemological	Structuring relations	Arbitrary	Structure
Ontological	**Ontological relations**	**Constrained** (meaning postulate s)	**Meaning**
Conceptual	Conceptual relations	Subjective	Conceptualization
Linguistic	Linguistic terms	Subjective	Language dependence

Entire disciplines exist for the definition and study of logical languages, which include an alphabet, with logical constants and non-logical symbols, a grammar and ways to study the interpretation and representation of truth, validity, and deductive statements.

[79] http://www.weblab.isti.cnr.it/talks/2009/iccu/images/thesVsOntology.png

5.2.1.1.1 Descriptor Logics

Some widespread logic languages such as formal deductive systems are First Order Logics and Descriptor Logics (DL).

Descriptor Logics are a subset of First Order Logics and are a widespread formalism to represent knowledge and semantic network. Reasoning is a central service, and inference is one of the primary characteristics of DL (Baader, Nutt, 2003). They are the base of some informatics and database languages sharing the same goals. In the built network, the *'is-a'* relation (that is the parent – children hierarchical relation) is essential, as it is in informatics ontologies developed starting from this basis. It is also useful to state constraints and perform inferences. The structure of DL is divided into two components: the 'TBox', which define terminology, and the 'ABox', which contains assertions. Several tools developed from the DL basis, permitted the representation systems to evolve and improve until the recent XML-based technologies for Semantic Web. DL generated further logical languages such as DAMN-ONT, OIL, DAMN-OIL, OWL, and so on (Nardi, Brachman, 2003). These are the base for the specification of the ontologies.

Many possibilities are given by the employments of descriptor logics, including the management of time, probabilistic logic and fuzzy logic. For details in their description see the bibliographic references (Baader, Nutt, 2003; Nardi, Brachman, 2003; Keller, 2004; Borgida, Serafini, 2003).

5.2.2 Ontologies: definition

In Computer Science, an ontology (with lowercase letter) is an information object, or computational artefact. Contrary to the philosophic idea, here the reality, what 'exists', is that which can be represented. It considers entities and relations useful to our purposes (Guarino, 2009).

Entities (the most general being) are analysed and organized into concepts and relations. The backbone of an ontology consists of a generalization/specialisation hierarchy of concepts, that is, a taxonomy (Guarino, 2009). This already makes clear the necessity of the employment of systems that implement hierarchies and inheritance in the final application phases.

In other words, (Smith, 2003) defines the ontology, using the words of (Quine, 1953): *'a network of claims, derived from natural sciences, about what exists, coupled with the attempt to establish what types of entities are most basic'*.

Some of the definitions of "ontology" are considered as the most meaningful, and are often cited in the reference bibliography.

'Explicit specification of a conceptualization' (Gruber, 1993)

'Formal specification of a shared conceptualisation' (Borst, 1997)

'An ontology is a formal, explicit specification of a shared conceptualisation' (Studer et al., 1998)

The last one merge the others. The meaning of the employed words is detailed in the following definitions.

'**CONCEPTUALISATION**' is an abstract, simplified view of the world, the portion of the universe of discourse that we wish to represent for some purpose. This is mathematically represented as a tuple $(D, \mathcal{R})$ where D is the universe of discourse and $\mathcal{R}$ a set of relations on D. More complex and articulated definitions can be given (Guarino, 2009). However, for our goals it is sufficient to consider that a conceptualization of a domain represents the invariants within and across presentation patterns. Ontological formulas are (assumed to be) the necessary information, that does not vary when changing representation.

'EXPLICIT SPECIFICATION' is the explicit communication of the intended conceptualisation, since it exists even if it rests in our mind. The exigency is to define a correct interpretation of the conceptualisation, since it is still unconstrained, and multiple interpretations are possible. An example of the consequences can be found in any comedy of misunderstanding.

The explication can be extensional (by listing every conceptual relation for all possible worlds). This cannot be done in an effective communication; at least some example can be usefully associated to some intensional explicating. For this last case, a language must be fixed and the constraints of interpretations are defined in that language by means of axioms (meaning postulates). The ontology is the set of these axioms.

'**SHARED'**: what can be shared are approximations of conceptualizations (meaning postulates) based on examples and showing the actual circumstances where a certain conceptual relation holds. The necessity rise from the fact that for an ontology definition one must always consider the limitation that its stakeholders must understand the primitive terms in the appropriate way. Therefore, for the ontologies that should support large-scale interoperability, it is important to be well-founded, with the basic primitives being sufficiently well-chosen and axiomatized to be generally understood (Guarino, 2009). This can be verified only through the sharing of them and their optimization starting from the users' feedback.

'**FORMAL'**: the used language must be machine-readable. This is the objective of a logic. This definition is presented as the last one since it implies the description of some linguistics and logic concepts requiring more attention to understand the use of ontologies in Computer Science applications.

'Formal' has two meanings. 'Formal logic' describes the connections between truths (neutral with respect to truth); 'Formal ontology' is the connections between things (neutral with respect to reality). For the reasons explained in the following, 'represented in a formal language' is not sufficient for being formal in this sense. Formal Ontology is the theory of formal distinctions and connections within entities of the world, as we perceive it (particulars) and categories we employ to talk about such entities (universals).

5.2.3 Ontologies in informatics

From the 1980s, some projects have been developed for the use of ontologies in computer science for applications in certain fields (such as medicines, business and so on). Some of them are listed in (Smith, 2003). Of a particular relevance is the CYC project, whose results are still integrated by more recently published languages for the development of Semantic Web[80].

The ontologies can be classified into different types, depending on the broadness of the considered domain (Guarino, 2009).

The '***top-level ontologies'*** (or upper-level ontologies) provide a broad view of the world, suitable for different target domains; the advantage in their use is apparent for information science applications: databases can be calibrated in terms of one common ontology (a single, consistent, stable, and highly expressive set of category labels) (Smith, 2003). They specify highly general (domain-independent) categories (time, space, inherence, instantiation, identity, measure, quantity, process, event, attribute, etc.). This is a solution to the problem of the trade-off between the fact that an ontology must be neutral to be maximally accepted, but it also needs to be expressively powerful and wide-ranging (containing therefore the largest possible number of terms) (Smith, 2003).

Some example of upper level ontologies are BFO[81] (Basic Formal Ontology); SUMO[82] (Suggested Upper Merged Ontology); DOLCE[83] (Descriptive Ontology for Linguistic and Cognitive Engineering); UFO (Unified Foundational Ontology) (Guizzardi, Wagner, 2004); others can be found at http://www.ontology4.us/Ontologies/Upper-Ontologies/Dolce%20Ontology/index.html.

The ***'reference ontologies'*** provide a reference for structuring ontologies;

The ***'core ontologies'*** define a super-domain, for example, the entire world of cultural heritage;

The ***'application ontologies'*** can be directly used in reasoning engines or software packages.

Systems implementing ontologies are, for example: KYOTO[84] (Knowledge Yielding Ontologies for Transition-based Organization) (Vossen et al, 2008); WordNet(s) (Fellbaum, 1998), and others.

5.2.3.1 The modelling process of an ontology

Following the phases of database modelling previously examined (Guizzardi, 2010):

- in conceptual modelling, a solution-independent specification is produced for a clear and precise description of the domain elements;
- in the design phase, the conceptual specification is transformed into a logical design specification considering different issues. The same conceptual specification can produce a number of different logical designs;
- in the implementation phase, a language is used to be deployed in a computational environment.

The same steps are followed in ontology engineering for modelling a domain-ontology.

---

80 http://sw.opencyc.org/
81 http://ifomis.uni-saarland.de/bfo/
82 http://www.adampease.org/OP/
83 https://en.wikipedia.org/wiki/Descriptive_Ontology_for_Linguistic_and_Cognitive_Engineering
84 http://weblab.iit.cnr.it/kyoto/xmlgroup.iit.cnr.it/kyoto/index8d0f.html?option=com_content&view=article&id=390&Itemid=155

In the conceptual modelling phase, the main requirement is the domain appropriateness (comprehensibility appropriateness). The models should be truthful to the phenomena being represented, and it should be clear for users to understand which real-world elements are represented by model elements and what problem-solving operations are to be performed. They must be computationally independent, in order not to be tied to an evolving technology.

Very often, UML is used for database conceptual modelling (to be even considered a *de facto* standard). Since it is considered by (Guizzardi, 2010) as an epistemological-level language, it requires some improvement for additionally being effective in modelling ontological conceptualizations. A proposal is described in (Guizzardi, Wagner, 2010), that use the UFO (Unified Foundational Ontology) for evaluating and redesigning UML. Similar tools are also used to analyse the relationships from an ontological perspective (Guizzardi, Wagner, 2008) and the representation of subsetting, specialization and redefinition relations (Costal et al., 2011).

The needs of ontology engineering for building effective ontologies are exposed in (Guizzardi, 2010), where some solutions, such as the use of design patterns for building ontologies and more have been proposed.

The phase of ontological design provides methodological supports for the exploration of solution space (supporting reasoning and satisfying non-functional requirements particular to a specific computational environment), and for mapping from the conceptual to the implementation level, preserving real-world semantics.

The ontology can be used to produce several different implementations in different codification languages (Ontology Web Language - OWL, Descriptor Logics - DL, Resource Description Framework - RDF, etc.). The choice of one of the possible codification languages will be justified as a design choice.

For validating ontologies, some methods were proposed, with the aim to describe the semantics of the ontology with more precision to check for potential inconsistencies in the taxonomic structure. Two of these are the OntoClean methodology proposed by (Guarino and Welty, 2002) and the OntoUML validator, developed by (Benevides, 2010) on the base of OntoUML methodology (Guizzardi et al., 2011).

5.2.4 Ontologies vs data model

A fundamental difference between an ontology and a conceptual model exists. An ontology is an open schema: one cannot suppose what is not told in the model (e.g. if it is not stated whether a capital has volutes, it can have them or not; we cannot know which is true) but we can extend the ontology in order to include more information.

Instead, in a conceptual model, only what is told is true (e.g. if a capital is not stated to have volutes, it does not have them). It is a closed model.

Furthermore, an ontology aims at representing a whole domain in the most generic and task-independent way possible, while a data model is intended for a specific application and it is not aimed to be shared by different applications (Spyns et al., 2002).

Sometimes some schema appears to have characteristics of both categories, but some character usually prevails and it can be critically considered in relation to the aspects of interest.

5.2.5 Domain ontologies in the modelling process of an information system

Generally speaking, ontologies play a double role in information technologies as a key structure for both database interoperability and information retrieval.[85]

Every Information System (IS) has its own ontology, since it ascribes meaning to the symbols used according to a particular view of the world. The peculiar role an explicit ontology can play within an information system is a central one, and the ontology profitably 'drives' all aspects and all components of an IS, so that we can speak of ontology-driven information systems. The most obvious use of an ontology is in connection with the database component. In fact, an ontology can be compared with the schema component of a database (Guarino, 1998). This should be primarily considered for structuring the intended system.

Fundamental aims of ontologies are (Guizzardi, 2010):

1) *semantic interoperability*, supported by their use as a reference model of consensus;
2) *automated reasoning*, possibly due to the use of ontologies as an explicit, declarative and machine-understandable artefact coding a domain model.

Even if automated reasoning is probably the reason for which the ontology technologies were originally developed, semantical interoperability is central to the necessities of database interoperability and information retrieval in the Semantic Web.

Requirements for realizing semantic interoperability are: to maximize the expressivity in capturing fundamental aspects of the underlying domain (and making the ontological commitments explicit); to maximize conceptual clarity to afford communication, domain understanding, problem-solving and meaning negotiation among human users. An ontology can be considered as a conceptualization method: to replace the domain of semantic interpretation (conceptualization) by an ontology (Laurini, 2012).

Moreover, they should serve as a support for reasoning in the Semantic Web (Guizzardi, 2010).

A domain ontology can be considered as a special kind of conceptual model that presents in addition a model of consensus within a community, to share information by conforming to a standard set of constructs. Furthermore, it can include constraints, business rules, derived rules, etc. (Laurini, 2012). In the coding process of an ontology, the real-world semantics should be preserved. The role of a domain reference model is to provide a reference frame, that is, to serve as a conceptual tool to master the complexity and to harmonise heterogeneous viewpoints and terminologies.

The flow of the database modelling should include the passage to an ontology (which should be standardised)-compliant conceptual model before to pass to the design phase (Figure 67).

The schema including a representation of the semiotic triangle for underlying the relationship between reality and the final model is shown in Figure 66 and described in the following example (Table 2).

[85] The italic text cites (Laurini, 2015)

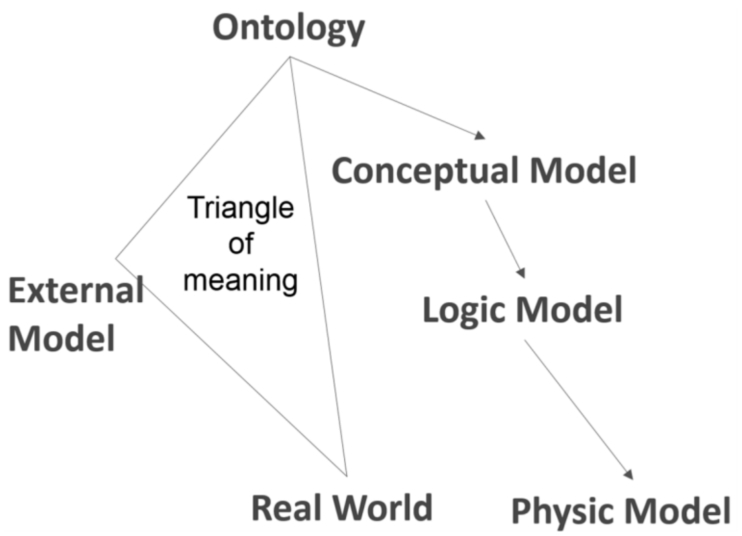


Figure 66. Database modelling workflow considering the triangle of meaning as reference to reality: the real world corresponds to the single objects; the description in an external model can be considered as the symbol level (which can vary for each person or culture or point of view); the abstraction to the unique concept level is formalized by the ontology. The ontology is the starting point for extracting a conceptual model, depending on the required application, and the following modelling phases of a database.

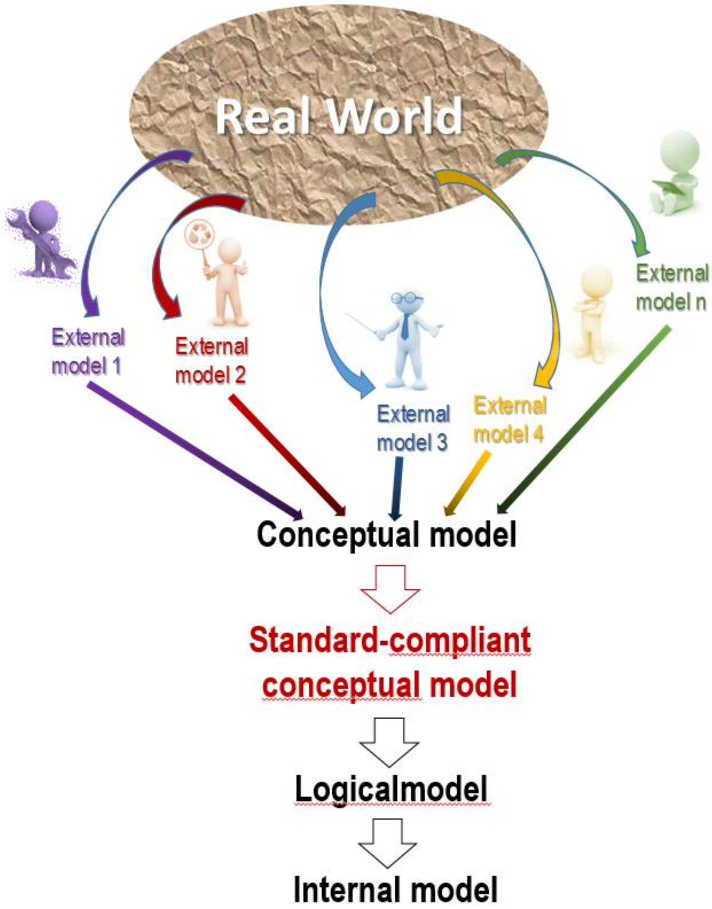
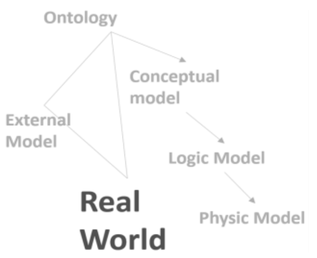
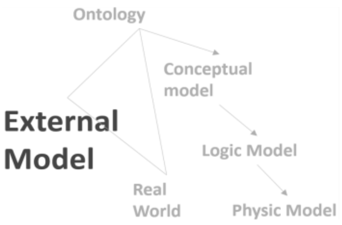


Figure 67. The ANSI/SPARC database modelling schema with the inclusion of the standardization of the conceptual model, to avoid the multiplicity of spatial representations and the lack of interoperability adapted from (Armstong, Densham, 1990 in: Laurini, Thompson, 1992).

Table 2 An example of the modelling process for a capital representation

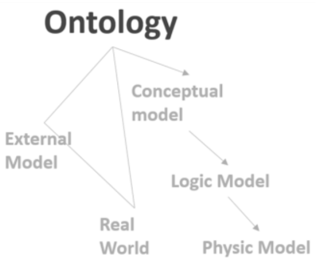


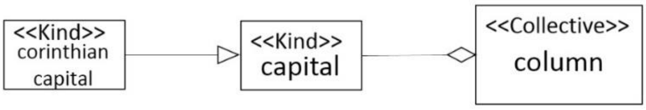
(a capital in the real world)



Ex. string
«c-a-p-i-t-e-l-l-o» (it)
«c-a-p-i-t-a-l» (eng)
…

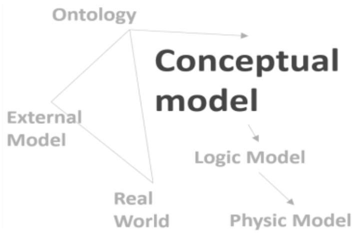
(the symbol used to comunicate the concept of similar real objects)

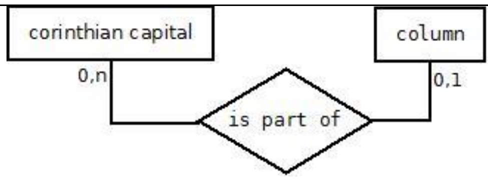




(description using an ontological language)

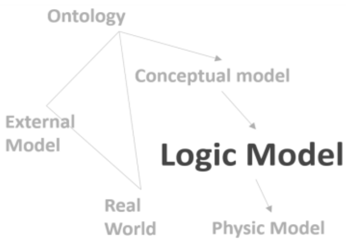
Following OntoUML rules: the kind (that is, a substance with identity) ‘corinthian capital’ is-a subclass of (hierarchical relation) the kind ‘capital’, which is-part-of (mereological relation) a collective (that is a whole) ‘column’.





(description using a formal epistemological language)

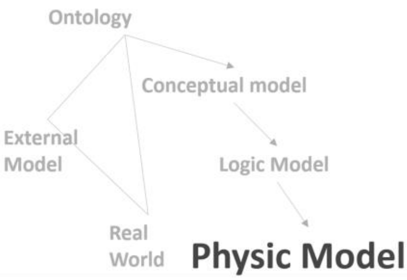
Example in Entity - Relation diagram

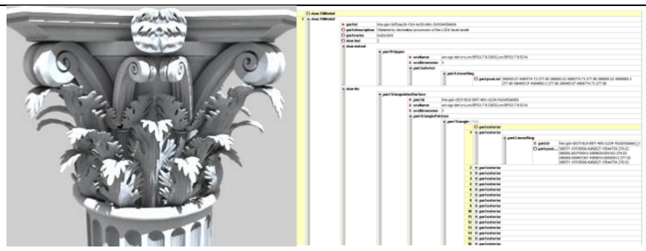



```
Es.
<rdf:Description rdf:ID=«ParteDiColonna">
          <rdf:type
            rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
 </rdf:Description>

<rdf:Description rdf:ID=«Capitello">
         <rdf:type
           rdf:resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
<rdfs:subClassOf rdf:resource="#ParteDiColonna"/>
</rdf:Description>
```

(description in an implementation-oriented language, e.g.RDF)





(implementation)

5.2.6 Semantic integration

(Laurini, 2012) affirms that originally in information technology, ontologies were created to solve interoperability problems between databases; in this context, each database must have its own ontology (usually a sort of re-writing its conceptual model) called 'local ontology'. For each concept of the local ontology, one or more concepts of the domain ontology must exist. A special program called a 'mediator' should undertake the translation for performing the integration among different DBs (Figure 68).

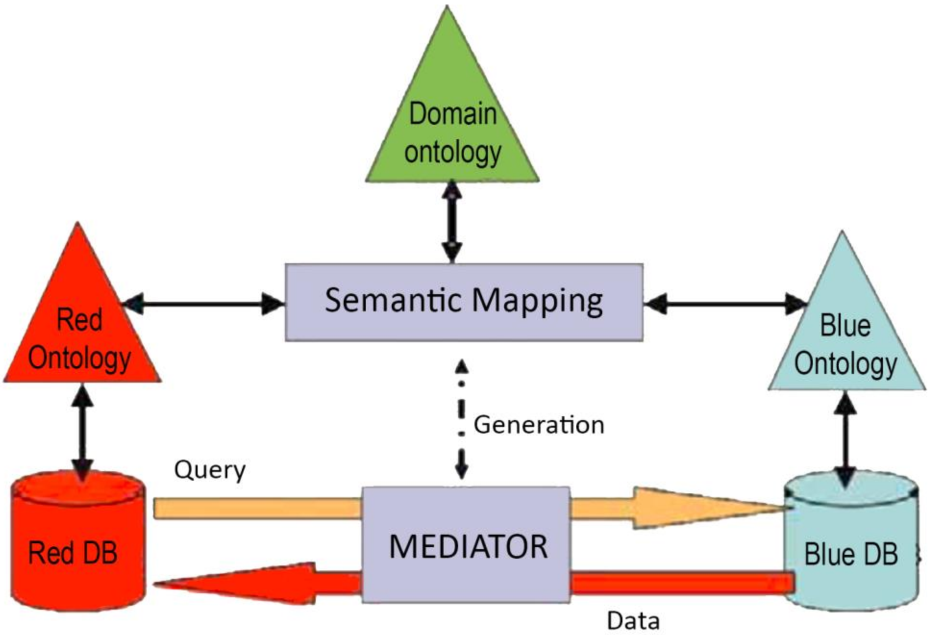


Figure 68. Translation and adaptation of the query and of the results by means of a mediator in charge of those transformations, driven by a domain ontology. reported from (Laurini, 2012).

The ontologies refer to general basic categories and relations (primitives). This permits the obtaining of some integration among databases of different applications that can be linked and compared, avoiding a 'data-sylos' effect.

5.2.7 Advantages of ontologies in information management and GIS

The original aim of ontology is reasoning about reality, and computers largely empower this possibility. The systems reviewed below are used for automatic reasoning (Ghidini, Giunchiglia, 1998).

In some research investigations, ontologies were used for structuring GIS. The aim was to exploit artificial intelligence issues and interoperability topics (Fonseca et al., 2000; Fonseca, Egenhofer, 1999). The main considered issues are the necessity to represent bona-fide and fiat objects and the necessity given by the use of ontologies to use object-oriented structures and multiple inheritance. In the investigations, some implementation issues were discussed. In this context, the use of object-oriented approaches was central.

Some urban ontology was developed (Fonseca et al., 2000; Keita et al., 2004), and further scenarios for ontology-based automatic processes in cartography were studied (Iosifescu-Enescu, Hurni, 2007).

For managing urban objects, it was clear that the inclusion of mereological concepts in the ontology was important, or, better, it was an effective use of a "mereotopology": an alliance of topological methods with the ontological theory of mereology. Other issues for the urban objects structuring and representation were the necessity of hierarchical structures and the possibility to combine multidisciplines' conceptualizations. The effectiveness of ontologies for these aims is clear.

However, after some experimentation with ontological GIS (Goodchild, 1997; Egenhofer, 1999, Fonseca et al., 2000; Fonseca, Mark et al., 2000; Mennis, 2003) the field was subsequently minimally investigated and spread. This was possibly due to the fact that technology had to develop in order to improve computation and storage possibilities. Moreover, the relational databases offered more immediate advantages, possibly more apparent to non-expert users. However, the currently available technologies can permit the substantial exploitation of object-oriented and ontology technologies. The economic dynamics cited when speaking about databases probably influenced this fact.

5.2.8 An example of ontological language: OntoUML

The meta-model OntoUML was proposed by (Guizzardi, 2005) as an ontological-level language (Guizzardi, 2010). It contains elements that represent ontological distinctions prescribed by an underlying foundational ontology and constrains that govern their relations. It offers some basic elements for building the ontology and structures to validate it. An automatic interface to be used for building and validation of ontological structures using OntoUML is implemented by a model-based graphical editor in the Eclipse environment developed by (Benevides, 2010; Benevides et al., 2010).

In an OntoUML metamodel, classes and relations are predefined to be support in modelling, since they are supposed to include all of the possible sorts of elements to be represented in the ontology. Some rules and constraints are then stated for validating the ontology.

5.2.8.1 A study of the application of OntoUML structures to cultural heritage

A study about intangible cultural heritage was realized by using the OntoUML graphical editor for the conceptual modelling of a case study: the Piedmont gypsum ceilings (Spanò, 2013), which are relevant to cultural heritage documentation since they carry some intangible values, such as the know-how that permitted their realization and the symbols and iconographies impressed on their surfaces.

The considered OntoUML classes have the following stereotype:

- 'Kind', the substance sortal that supplies a principle of identity and congregates all the essential properties of its instance;
- 'Role', phase-sortal that, in contrast, represents the contingent properties of an object (kind) in a relational context; it is an anti-rigid and relationally dependent universal.
- 'Relator', individuals with the power of connecting entities
- 'Mode': intrinsic moment universal. Every instance of mode universal is existentially dependent on exactly one entity. Examples include skills, thoughts, beliefs, intentions, symptoms, and private goals.

The meta-relations here considered are:

- 'Formal' hold directly between two entities;
- 'Material' induced by mediating entities, the relators:
- 'Mediation' is the mediation relation holding between a universal U and a relator universal UR iff (if and only if) every instance of U is mediated by an instance of UR;
- 'Generalization' (subclass-superclass relation);
- 'Component of' (object part-of another object);
- 'Member of' (object part of a collection of objects);
- 'Relation Derivation' represents the formal relation of derivation between a material relation and the relator universal this material relation is derived from.

These classes and the constraints to which they are subject are implemented in a graphical interface using the Eclipse-based frameworks (Benevides, 2010). This permits the automatic verification of the validity of the built model.

A reduced model of the considered topic has been represented focusing on the nature of the entity representing the intangible part of cultural heritage (Figure 69). This core model is extended by adding other elements that specify the topic and could be useful to the representation. These are the kinds 'Craft', 'Decoration', 'Building', 'Municipality', and 'Value' that could, in turn, be more specified.

The considered object is the kind '*PlasterCeiling'* that served as was our concrete case study. It becomes a role '*CulturalHeritagePlasterCeiling'* by virtue of its material relation with some '*Civilisation Value'* that we can identify through two chosen relator: the *'Know-how*' of artisans who produced that hand-made object and some recurring '*Symbols*' reproduced on the decorations. These last two entities represent the intangible cultural values linked to the material cultural heritage, and it could be meaningful to highlight that it is this relation that makes some common object part of a cultural heritage.

If we check the validity of this model, through the specific tool implemented in the Eclipse open application environment, we find that mereological relations ('componentOf' and 'memberOf') are invalid (see the red symbols on the relations in Figure 69). That is because the imposed constraints state that a whole must have at least two disjoint parts. This does not affect the validity of our model, as we did not represent the remaining parts for simplifying the model, but we know them to exist.

Some problems could instead derive from the invalidity of the material relation between '*CulturalHeritagePlasterCeilings'* and '*CivilizationValue'*. Here the problem is that the relation is mediated by two alternative relators, while the rule states that there must be only one relator for each material relation. For such a complex thing as cultural heritage is, this rule could be restrictive. It is common knowledge that a cultural heritage object has a number of characteristics meaningful for the determination of its value and for which it has to be preserved and documented. Some of these values are listed in the selection criteria for UNESCO Cultural Heritage (UNESCO, 2013), in the UNESCO Convention (UNESCO, 1972), and in many official and unofficial documents connected to cultural heritage. It could be of importance to consider in the model the possibility of all these values, both as alternatives or as present at the same time. The same constrain problem affects the cardinality of the mediation relation, which can only be set with the lower value '1', making the relation obligatory. This is not true if, for some instance, one considers only one relator (such as the know-how used for its production; this is the case of some ceilings found in Sicily).

Through this auto-validated model, we could outline the conceptual model of our topic defining which entities we need and how they relate to each-other. Moreover, from this model, the instances of intangible cultural heritage present in our case study clearly emerge. Additional ones can be inferred in extensive datasets.

The ‘*know-how*’ is formally an Intangible Cultural Heritage (ICH), as stated by the UNESCO Convention for Safeguarding of Intangible Cultural Heritage (2003) (UNESCO, 2003), while the represented symbols could be associated to some folkloric tradition of people living in a place, but there does not exist a specific categorization describing them.

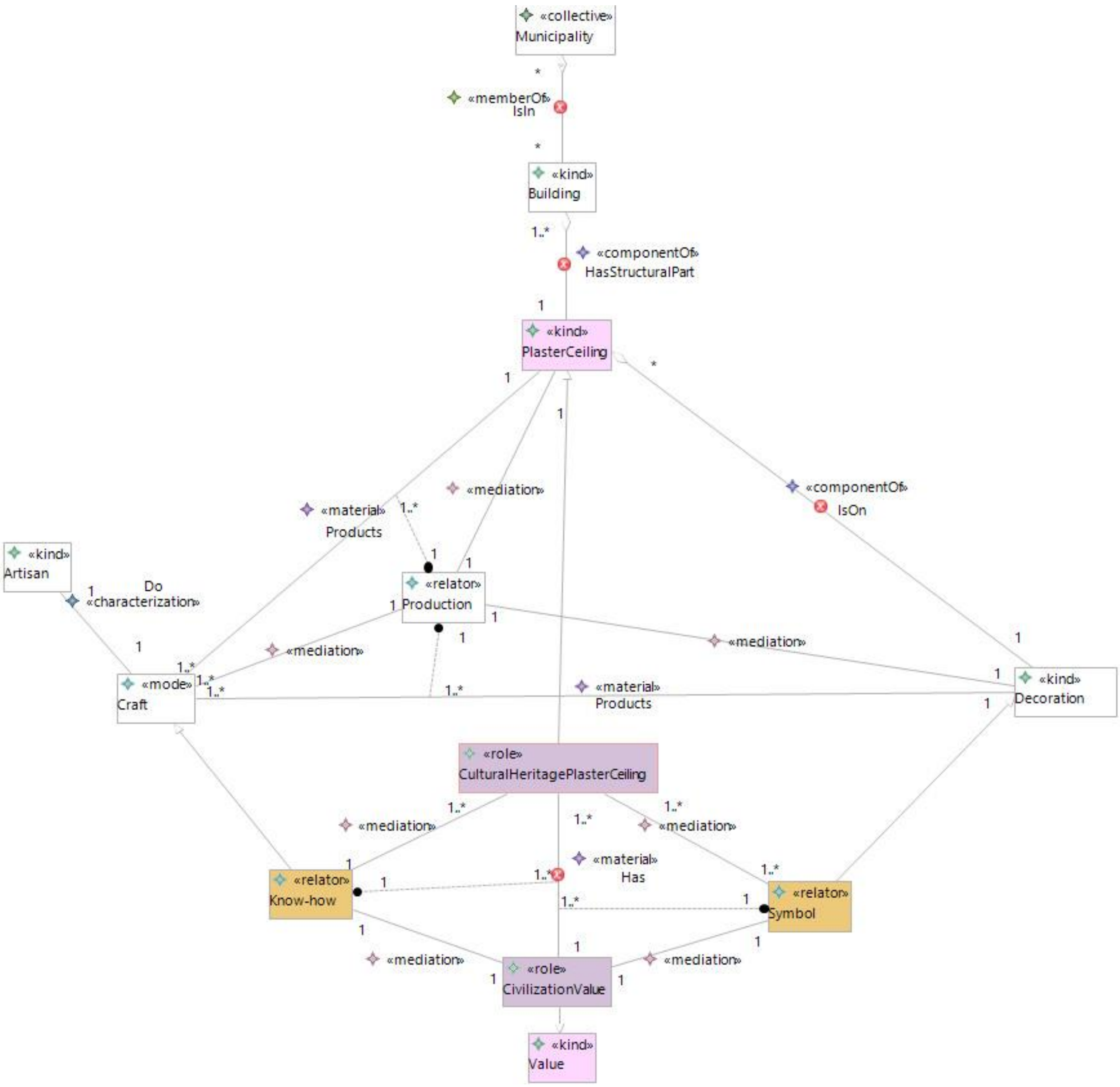


Figure 69. OntoUML model processed using the eclipse-based graphical interface. In yellow are the entities which are Intangible Cultural Heritage items.

6 INTEROPERABILITY EXIGENCIES: STANDARDISATION

"In recent years there has been a growing recognition of the benefits of creating cultural heritage information networks that will enable common access to documentation created and managed by diverse organisations. An important precondition to the development of such networks is the documentation standards that establish the degree of compatibility needed to make common access possible. […]"

Eleanor E. Fink
Director, The Getty Information Institute[86]

The interoperability and the harmonization of the data for common interpretation and reuse requires some common language to be used. This language can be only defined by the community, which must choose and adopt some common data model for organizing the archived information. With this aim, several organizations, of different natures, are collaborating.

Standards are recognised as one of the foundations of the Network, as they should be for any scientific activity. Cooperation and integration are indeed impossible, if all the interested parties do not agree on the way in which data are stored, exchanged and output for further processing. The pipeline concept implies that outcomes of a process are inputs to the next one, what is impossible if there is no match between the ways output is produced by the former and input are expected by the latter. [87]

The foremost aim of international standardization is to facilitate the exchange of goods and services through the elimination of technical barriers to trade.[88]

In various fields of information management, specific standards have been defined. Cultural heritage and cartography or spatial data management fields are herein considered.

From the cultural heritage point of view, information sharing is important for several reasons, including the promotion and interpretation of the heritage for vital economic purposes (such as cultural tourism and regional development) and the reinforcement of cultural and social identity at regional, national, and international levels[89].

Although documentation of the cultural heritage is already carried out at local and national levels, the need to use information produced by documentation centres is becoming international in scale, responsive to the global trends in economic activity and the increasing communication potentialities of the internet. If effective standards are available, considering an international perspective, the information can be shared and can constitute a powerful network.

However, the risk in cultural heritage standardisation is the loss of detail in description, which permit the inclusion in the documentation all the characteristics that cause a cultural heritage item to be a cultural heritage item. For this reason, it is important not to forget the data that permit the description of the main characteristics of a cultural object, including its material persistency, shape, and dimensionals characteristics, as well as historical facts. Moreover, the information about the management of the object (state of conservation, ownership, and so on) are important to its preservation. Finally, the multifaceted information and nuances must be included in the documentation in order to effectively communicate the essence of the object. In the case of architecture the location

[86] The italic text cites http://archives.icom.museum/objectid/heritage/getty.html
[87] The italic text cites http://epoch-net.org/site/research/standards/
[88] The italic text cites (ISO, 2015).
[89] http://archives.icom.museum/objectid/heritage/intro2.html

of the monument in its context, the distribution of spaces, the internal and external paths, the urban function and further characteristics are also fundamental.

This kind of information is the subject of further disciplines dealing with cartography and 3D modelling. A corpus of standards and laws is available in this field. The developed directives are also essential for different application fields, such as environmental management and risk prevention, or for political reasons. Therefore, the available standards are often oriented to the management of such kinds of information and analysis. Another limiting issue of the cartography standard is the consideration of a representation scale that is insufficient to effectively describe architectural heritage.

6.1 International Organization for Standardization (ISO)

Different organisations and institutions join their efforts to redact and publish standards in various disciplines, but one of them deserves particular attention. This is the ISO, which is the reference for affirming the international dignity of standards. ISO is the international body responsible for the standardisation of almost all sectors.

ISO is a legal association, the members of which are the National Standards Bodies (NSBs) of some 164 countries (organizations representing social and economic interests at the international level), supported by a Central Secretariat based in Geneva, Switzerland.

The principal deliverable of ISO is the International Standard.

An International Standard embodies the essential principles of global openness and transparency, consensus and technical coherence. These are safeguarded through its development in an ISO Technical Committee (ISO/TC), representative of all interested parties, supported by a public comment phase (the ISO Technical Enquiry).

Geographic information standards are recognized as an underpinning in the realization of SDIs (Spatial Data Infrastructures). They have been developed extensively since 1994 for the sharing and the integration of geographic data across platforms. They facilitate the discovery, the access, and the appropriate use of geographic information, i.e. interoperability. Interoperability of geographic information is the essential goal of geographic information standards. It also relies Information technology (IT) standards along with domain specific knowledge standards. They are providing all together the pillars for interoperability (Figure 70). Standards provide consistent and interoperable definitions and structure for data and metadata (semantics), encoding (syntax) and access (services).[90]

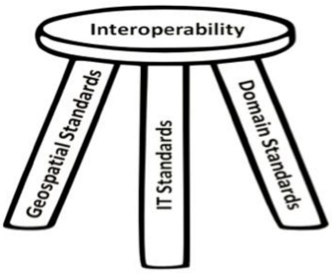


Figure 70. The standards pillars (ISO, 2015)

In this chapter, some relevant existing standards for cultural heritage (CH) and spatial representation are presented.

[90] Italic text cites (ISO, 2015).

6.2 Cultural Heritage Standards

Article 17 of the 1985 Granada Convention requires parties to exchange information on "the possibilities afforded by new technologies for identifying and recording the architectural heritage." Taking its cue from the Convention, a roundtable was convened in London in 1989 to examine the tasks of architectural heritage information centres, the ways and means of improving cooperation between them, and the new technologies available to them in furthering their work. Among the recommendations of the meeting was the following:

The standards relating to a minimum set of data elements and the technical specifications required for their communication should be identified. This should be done by determining which data elements are necessary for the recording of all buildings of historic and architectural interest in each state or institution for the furtherance of its own work; by determining how this data may be harmonised; and by setting standards for computer systems. [91]

A number of local, national, or international standards affecting cultural heritage documentation and representation exist. The link of new generated information to these existing databases, inventories, or vocabularies should be permitted by new systems in order not to lose, on the one hand, and to enrich, on the other hand, the archived information. Some of the standards are reported and briefly described in the following subchapters. They follow an approximately chronological order.

6.2.1 Heritage Documentation Programs

USA - from 1933

The Heritage Documentation Programs (HDP) are part of the American National Park Service, which provides some standards and guidelines for the documentation of American cultural heritage. The standards consist essentially in some principles to be respected when documenting monuments (they are published in the Federal Register on September 29, 1983 Vol. 48, No. 190, pp. 44730-34). The guidelines provide some requirements about how the documentation must be acquired and presented and they recommend a list of attributes to be filled in when documenting cultural heritage. They were originally published in the Federal Register on September 29, 1983. A revised version was published in the Federal Register on July 21, 2003 (Vol. 68, No. 139, pp. 43159-43162)[92]. The HABS guidelines (Historic American Building Survey) are the Federal Government's oldest preservation programme (begun in 1933 to document America's architectural heritage). They are divided into history guidelines[93], which mainly include the thematic data to be documented and drawing guidelines[94], which suggest the way to execute the survey drawings. Finally, photography guidelines[95] provide indications on what material and techniques to consider when taking and processing documentation photographs.

[91] The italic text cites http://archives.icom.museum/objectid/heritage/intro3.html
[92] http://www.nps.gov/hdp/standards/index.htm
[93] http://www.nps.gov/hdp/standards/HABS/HABSHistoryGuidelines.pdf
[94] http://www.nps.gov/hdp/standards/HABS/ HABSDrawings.pdf
[95] http://www.nps.gov/hdp/standards/PhotoGuidelines.pdf

Heritage Documentation Programs administers the Historic American Buildings Survey (HABS), the Federal Government's oldest preservation program, and its companion programs: the Historic American Engineering Record (HAER) and Historic American Landscapes Survey (HALS). Documentation produced through the programs constitutes the nation's largest archive of historic architectural, engineering, and landscape documentation. The HABS/HAER/HALS Collection is housed at the Library of Congress.[96]

HDP also develops and maintains the Secretary of the Interior's Standards and Guidelines for Architectural and Engineering Documentation.

HDP conducts a nationwide documentation program in partnership with state and local governments, private industry, professional societies, universities, preservation groups, and other Federal agencies. The program assigns highest priority to sites that are in danger of demolition or loss by neglect, and to National Park Service properties. In addition to the summer recording program, documentation enters the Collection through mitigation activities under appropriate sections of the National Historic Preservation Act of 1966, submissions in prize competitions, and donations.

Documentation provides a permanent record of the nation's most important historic sites and large-scale objects. The Collection is unique in the strong support it enjoys from its institutional sponsors and the public, and is distinguished in its national scope, consistent format, archival stability, and continued growth.[97]

HABS recording combines drawings, history, and photography to produce a comprehensive, interdisciplinary record. The documentation ranges in scope depending largely upon the level of significance and complexity. It should first and foremost convey what is most important about that particular structure.[98]

6.2.2 NORMAL

Italy – from 1977

A standard that is not directed to the general documentation of cultural heritage but to the definition and mapping of pathologies and degradation of monuments was started in Italy by the Commissione NorMaL (NORmalizzazione MAteriali Lapidei) headed by Giovanni Urbani (ICR Director). The aim of the Commission was to define a unique standard methodology to study stone materials (both natural and artificial, such as clay) and their alteration, in order to control the preservation, intervention, and treatments on historic and artistic artefacts, in compliance with the Italian restoration and preservation philosophy. The proposal was first presented in 1977 in the International Symposium on Conservation on Stone Materials, to the International Institute of Conservation (IIC), with a presentation having the title ‘Artistic Stone Works - A proposal for the Unification of the Methods of Studying Stone Decay and of Controlling Stone Conservation’. Different operators participated in the redaction of the first recommendations, under the coordination of the Consiglio Nazionale delle Ricerche (CNR) and the Italian Cultural Heritage Ministry (Ministero dei Beni Culturali).

From 1996 the Commission structure underwent a modification, with the convention between the Italian Cultural Heritage Ministry and the Italian Organization for Standardization (UNI). The common goal was to establish national technical rules to be proposed in Europe for the constitution of

[96] The italic text cites http://www.nps.gov/hdp/
[97] The italic text cites http://www.nps.gov/hdp/about.htm
[98] The italic text cites http://www.nps.gov/hdp/standards/habsguidelines.htm

a harmonic body of laws for the restoration. To this new Commission, the European Committee for Standardisation (CEN) could participate.

6.2.3 ICOMOS-ISCS Illustrated glossary on stone deterioration patterns

International Community – 2008

The ICOMOS International Scientific Committee for Stone (ISCS) glossary is an important tool for the classification and terminology for managing stone deterioration by both researchers and other operators.

It starts from the terms and classification of several further documents, including the Italian standard NORMAL to build shared guidelines for intending the same meaning when speaking about stone and its decay (ICOMOS-ISCS, 2008).

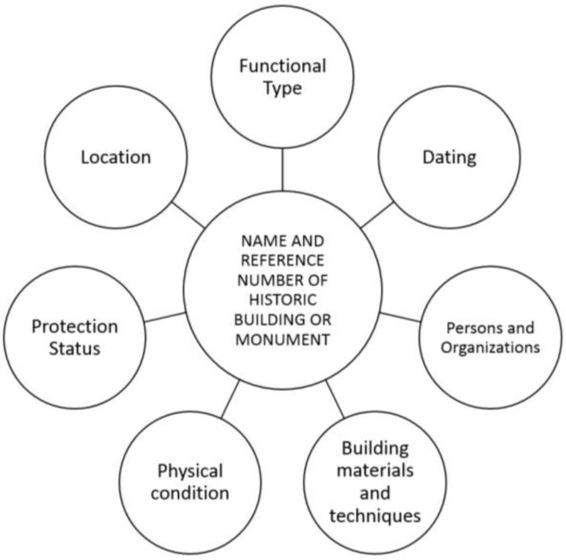
6.2.4 Core Data Index to Historic Buildings and Monuments of the Architectural Heritage

Europe - 1992

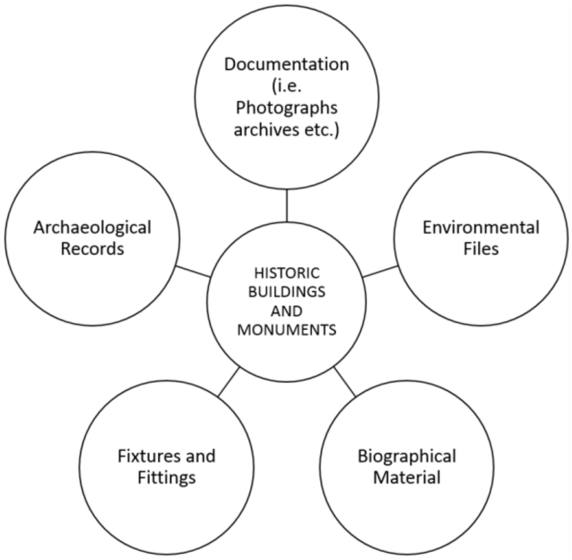
This standard was published in 1992 by the Council of Europe and the Getty Information Institute in order to foster the documentation and the standardisation of architectural monuments internationally, having acknowledged the importance of the standardisation aims. It was developed in parallel with two other core standards: the International Core Data Standard for Archaeological Sites and Monuments (1995), for documenting the immovable archaeological heritage, and Object ID (1997), to identify cultural movable objects.

The basic aim of the Core Data Index is to make it possible to classify individual buildings and sites by name, location, functional type, date, architect or patron, building materials and techniques, physical condition, and protection status. It is not an end in itself, but a starting point, a key to further information held in databases, documentation centres, and elsewhere that is necessary for the detailed understanding and care of individual monuments.

The Index is designed to enable the compiler to make cross references to the more detailed information about a building, including written descriptions and photographs; associated archaeological and environmental information; details of fixtures, fittings, and machinery installed within individual buildings; and the information on persons and organisations concerned with its history (Figure 71).



.I



.II

Figure 71. (I) The representation of a record structure for a building or monument and relationship of its parts. (II) Representation of potential relationship between core monument records and related more comprehensive levels of information.[99]

The Index has the potential not only to record individual buildings, but also to enable the compiler to relate the building to a larger site of which it may be a component or to the still larger ensemble of which it may form a part. The architectural ensemble manifests itself in many different forms. It may be typologically or geographically defined. It may be planned or organic, unified or grouped by association, or united by a common functional purpose or community of interest. It may be based on the hierarchical relationship between a larger structure and its components, such as apartments in a house or the machinery in a factory. It may be spatial, involving the considerations of the relationships between buildings, the spaces between them, and the landscape in which they sit. Different cases and organisational priorities will result in ensembles being defined in varying ways according to circumstance, imposing cut-off points in different places, in order to make the material manageable and to allow the making of connections that will permit a more rounded view of the heritage.

The Index does not seek to impose a rigid system, or to force organisations to act outside their own areas of interest. Nor does it seek to specify the computer hardware and software requirements of those organisations that are engaged in the process of computerising their information. Rather, it represents the first step towards defining and recommending technical standards for data capture and data exchange. It is possible to envisage a situation in which the mutual interrogation of indexed information will enhance our understanding of the architectural heritage of Europe. The Index is an important milestone on this road.[100]

A detail of the standard structure is available at http://archives.icom.museum/objectid/heritage/core.html, and a short history of the standard is explained at http://archives.icom.museum/objectid/heritage/intro3.html.

It is possible to note that a list of attributes is proposed for the documentation of the objects. It could be implemented using a relational DB structure. Also, the necessity of having a unique identifier of the resources is already clearly stated, hence the necessity of establishing relations between monuments and with further information for the great potential that this can offer. In the foreword parts of the document a useful documentation method in stated to be almost independent of the

[99] The italic text cites http://archives.icom.museum/objectid/heritage/intro3.html
[100] The italic text cites http://archives.icom.museum/objectid/heritage/intro3.html

management of the information using computers, exposing the application-independency principle, which is still clear for the definition of data models and standards.

The ICOM's International Committee for Documentation (CIDOC) is realizing a further combined standard for the inventory of both archaeological and architectural heritage: the 'International Core Data Standard for Archaeological and Architectural Heritage'[101].

6.2.5 Guarini archives

Regione Piemonte, Italy – from 1992

The Guarini project is funded by the Regione Piemonte, in Italy for improving the safeguarding, intervention, and promotion of cultural heritage by public or qualified institutions. The procedure acquired the dignity of a standard in 1994 with the decision of the Giunta Regionale (deliberazione n. 368-37612 del 3.8.1994). The established rules are expressed in some regional laws: the standard 'Patrimonio Culturale', the versions 'Censimento' (L.R. 34/95 and L.R. 35/95), 'Beni Librari' and the more recent 'Archivi'. The components will constitute the cultural heritage information system of the Regione Piemonte, to permit the integrated and extended data retrieval through the web of Piedmont cultural heritage. The modules follow a tree structure and are compliant with some further higher-level standards, such as the ISAD (International Standard of Archival Description), ISAAR (International Standards of Archival Autority Records) and the Italian standards defined by the ICCD. The geographic information as georeferenced entities is not included[102]

A specific software is also provided in order to respect the standard guidelines and a web browser for the catalogue is in development.

6.2.6 ICCD documentation forms

Italy – from 1998

The Italian Istituto Centrale per il Catalogo e la Documentazione (ICCD) is the institute dependent from the Italian Ministry of Cultural Heritage and Activities for the research and development of tools and methods for the documentation and knowledge of cultural heritage to foster its protection and safeguarding. It established some standard forms for the collection of data about cultural heritage items and the progress towards completion of a general catalogue of cultural heritage in Italy. The catalogue is still incomplete and can be consulted online[103]. It is a collection of unrelated forms that collect the data about cultural heritage objects, controlling and codifying the data acquisition with criteria and predefined attributes. A further tool consists of the authority files, which regard the entities that are not cultural heritage item themselves, but can be related to them (authors, sources…). These are independent parallel databases related to the monuments catalogue, but they permit some support to the standardization. Recently a new form has become available, which is unique and therefore has the advantage of being suitable for achieving a major interoperability (Mancinelli, 2015b). The cultural heritage catalogue is organized in the SIGECweb, a digital system that permits access to the system through the web, organizing the data in digital databases and adding the visualisation of the location

[101] The italic text cites http://www.getty.edu/conservation/our_projects/field_projects/arches/arches_overview.html
[102] http://www.regione.piemonte.it/cultura/guarini/index.htm
[103] http://www.catalogo.beniculturali.it/sigecSSU_FE/ Home.action?timestamp=1452186106974

of the documented monument on a web map[104]. Recent development of the SIGECweb includes the effort towards the translation of the archived data to Linked Open Data formats (Section 7.2) and their mapping to the CIDOC-CRM (Section 7.8.1) (Mancinelli, Negri, 2014).

A GIS application of the ICCD documentation forms is the 'Carta del rischio', which aims at the preservation of cultural heritage, by interpolating some risk factors, such as environmental vulnerability and territorial dangerousness for determining the degree of risk to which the monuments are exposed[105]. It is managed by the Istituto Superiore per la Conservazione (before the ICR Istituto Centrale per il Restauro) as a support for scientific and administrative institutions for the safeguarding, protection and preservation of immovable cultural heritage items. The base element consists of a GIS, a technical tool for producing thematic cartographic representations combined with alphanumeric data. The first GIS of the Risk Map was developed between 1992 and 1996[106].

6.2.7 European digital Cultural Heritage inventories

Europe – from 2000

Europe has been funding several projects for the digitization of several cultural heritage categories (the architecture is not included) and the connection of the resources in European networks of catalogues and institutions following the MINERVA project, from 2002[107]. Between the years 2000 and 2012 the European Commission developed a set of tools with the aim of building a network of institutions dealing with cultural heritage and digital documents about cultural heritage in Europe. The used data models anyway are not yet the ones defined by the standards that can interest architectural heritage. For example, the OAIS (Open Archival Information System) reference model is used. Moreover, may be the established models are considered themselves as possible standards for the specific sectors.

The projects and their results are interlinked and continue to be developed. In the following, some of them are cited, but an exhaustive documentation can be found on the internet. An interesting timeline of the development of the digital cultural heritage network can be seen at the web page http://www.michael-culture.eu/about.

6.2.7.1 Multilingual Inventory of Cultural Heritage in Europe (MICHAEL)

Devoted to European cultural heritage valorisation, Michael Culture Association gathers a strong network of more than 100 public and private organizations from all over Europe. Key actor in the promotion and valorisation of the digital cultural content, Michael Culture develops tools and services for cultural institutions and the general public. Linked to other major European cultural heritage networks and projects such as Europeana, Europe's Digital Library, Michael Culture Association supports European and national cultural policies. It also aims at enhancing the network of European professionals working on digital cultural heritage, through the actions towards the Minerva Network.[108]

[104] http://www.iccd.beniculturali.it/index.php?it/376/il-sistema
[105] http://www.cartadelrischio.it/
[106] http://www.cartadelrischio.it/eng/info.html
[107] http://www.minervaeurope.org/
[108] The italic text cites http://www.michael-culture.eu/

One of the important tools that was developed, and that is being updated and improved, is the MICHAEL European portal: www.michael-culture.org. Through the multilingual MICHAEL service, people are able to find and explore European digital cultural heritage material using the internet. The MICHAEL service provides access to digital collections from museums, libraries and archives.[109]

6.2.7.2 Europeana

Europeana is a project funded by the European Commission in order to build a virtual European library to make Europe's cultural heritage accessible to all. The prototype was implemented in 2008 to begin the collection of books, paintings, films, museum objects and archival records throughout Europe in a digital form[110].

Europeana Digital Service Infrastructure (DSI) - 2015-2016

Michael Culture is full partner in the Europeana DSI project, as the aggregator for museums. Europeana is developing the Europeana Digital Service Infrastructure (DSI). DSIs are composed of 'core service platforms' which enable trans-European connectivity and interoperability, and related 'generic services' which link national and sectorial infrastructures to the platforms. Under funding from the Connecting Europe Facility, Europeana will develop into a widely recognised platform of services and resources, not only for metadata references, but also for access to cultural content, tools and technologies, projects and other services. [111]

6.2.7.3 AthenaPlus

AthenaPlus is a CIP (European Competitiveness and Innovation Framework Programme) best practice network started in March 2013 and ending in August 2015. The consortium is composed by 40 partners from 21 Member States countries. AthenaPlus will build on the successful experience developed by the previous ATHENA project – where LIDO (Lightweight Information Describing Objects) and the ATHENA Ingestion Server and Mapping Tool (MINT), widely used across the Europeana's ecosystem of projects including the ongoing Linked Heritage project were developed, in order to further advance and complete the effective infrastructure and tools developed to support museums and other cultural institutions in their work to making available digital content through Europeana.[112]

[109] The italic text cites http://www.michael-culture.eu/michael-multilingual-inventory
[110] http://www.europeana.eu/portal/
[111] The italic text cites http://www.michael-culture.eu/
[112] The italic text cites http://www.michael-culture.eu/athena-plus-646

6.2.8 Cultura Italia and PICO (Portale Italiano della Cultura On-line)

Italy – from 2005

The Italian Ministry of Cultural Heritage and Activities published in 2005 the Italian online portal of the Culture (PICO)[113]. The aim of the portal is to communicate Italian culture aspects and events to various kind of users (specialized operators or unspecialized citizens). The advantage of the portal is the indexing of the contents so that the data can be related to further resources.

*Il progetto: il Portale è promosso dal Ministero per i Beni e le Attività Culturali, che ha affidato alla società Politecnico Innovazione nel 2002 uno studio di fattibilità relativo, alla Scuola Normale Superiore di Pisa l'incarico del progetto tecnico-*scientifico *e al dipartimento di Caratteri dell'Architettura, Valutazione e Ambiente (CAVEA) della Facoltà di Architettura dell'Università degli Studi di Roma "La Sapienza" l'incarico di definire funzionalità e soluzioni tecnologiche della piattaforma di georeferenziazione GIS.*[114]

The aim of that project is the integration of the several existing data sources for a unique management of the resources and their completion when they are not exhaustive. This harmonisation requires a basic schema for the modelling of the data. This has changed from the existing ICCD documentation forms (useful for the construction of a relational database) to the more complex schemas and standards published and spread at the international level, and this aspect must be obviously continuously updated following the evolution of international standards. The data and metadata are mapped to the specific national and international standards and encoding systems, using application profiles. Details of the mapping can be found in http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/documenti/progetto/sintesi_progettotecnicoscientifico.pdf.

Integrazione di banche dati

La progettazione del modello di organizzazione delle informazioni nel Portale è stata determinata dalla necessità che il Portale offra accesso integrato al più ampio insieme possibile di informazioni e documentazione riguardante archivi, musei, siti archeologici, monumenti, patrimonio immateriale, patrimonio culturale digitale che saranno rappresentati all'interno del Portale sotto forma di metadati.

L'utente ha la possibilità di interrogare l'insieme di questi metadati attraverso un unico sistema di ricerca, raggiungendo poi l'informazione originaria sulle singole banche dati, che resteranno quindi indipendenti e non saranno duplicate. In parte i data source sono già individuati ed esistenti, in quanto afferenti alle direzioni centrali del Ministero (SIGEC, SITIA, SITAP, Carta del Rischio dell'ICR, Direzione degli Archivi, etc.); in parte essi non sono ancora individuati, ma già esistenti (ad esempio i Portali Regionali); in parte, infine, saranno da realizzare in concerto con il Portale. Sta ai singoli fornitori definire il concetto di risorsa, scegliendo le informazioni che siano meglio rappresentative delle entità che vogliono far conoscere attraverso il Portale e proteggendo, se necessari o, le informazioni sensibili o riservate. Pertanto caratteristica essenziale del metadata schema del Portale

[113] www.culturaitalia.it

[114] The italic text cites http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/documenti/progetto/sintesi_progettotecnicoscientifico.pdf.

The project: the portal is promoted by the Ministry of Heritage and Culture, which has entrusted the Politecnico Innovazione company in 2002 with the relative feasibility study, the Scuola Normale Superiore in Pisa the task of the technical-scientific project and the Department of Characters Architecture, Evaluation and Environment (CAVEA) of the Faculty of Architecture, University of Rome "La Sapienza", the task of defining features and technological solutions of the geo-referenced GIS platform.

è che supporti il requisito della scalabiltà, cioè che l'ingresso progressivo di nuove tipologie di entità informative non pregiudichi il funzionamento dell'intero sistema.

L'Application Profile del Portale

Per compatibilità con le scelte operate in ambito internazionale, e in accordo con le linee guida indicate dal progetto MINERVA (MInisterial NEtwoRk for Valorising Activities in digitisation, eContent*plus* - Supporting the European Digital Library[115])*, l'importazione di metadati (ovvero le informazioni che descrivono le risorse) da data source esterni nel Portale avverrà con il mapping in un unico schema. Tale schema è espresso secondo lo standard Qualified Dublin Core, che è stato ulteriormente esteso per definire l'Application Profile progettato per il Portale. Il Qualified Dublin Core è una versione derivata del Simple Dublin Core, codificato dal Dublin Core Metadata Iniziative. Comprende un set di elementi di base con i quali è possibile descrivere ogni tipo di risorsa. Per risorsa si intende "qualsiasi cosa che abbia un'identità".*[116]

In the project, a thesaurus of Italian terms has also been published for the classification of resources coming from different data sources, the PICO Thesaurus[117]. It conforms to the interoperability standards and tools, such as the XML format and SKOS systems of classification[118] (Simple Knowledge Organization System, published by W3C). It has to be interfaced with the international vocabularies for the managed resources to be retrievable from other systems in the world, as well. This is possible using an application profile[119].

Since the cultural heritage standards are being updated and new references are being implemented and published, the system has to adapt and has to be mapped to them. Therefore, some mapping rules and documents are compiled and can be applied to the resources for performing effective queries. From the web page http://www.culturaitalia.it/opencms/documentazione_tecnica_it.jsp the mapping

---

115 http://www.minervaeurope.org/home.htm

116 The italic text cites http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/documenti/progetto/sintesi_progettotecnicoscientifico.pdf

Integration of databases

The design of the organization model of the information in the portal has been determined by the need that the portal provides integrated access to the widest possible set of information and documentation concerning archives, museums, archaeological sites, monuments, intangible heritage, and digital cultural heritage that will be represented at the Portal inside in the form of metadata.

The user has the possibility to query all these metadata through a unique research system, and finally reach the original information in individual databases, which will therefore remain independent and will not be duplicated. In part, the data sources are already identified and existing, as pertaining to the central offices of the Ministry (SIGEC, SITIA, SITAP, ICR Risk Map, Directorate of Archives, etc.); in part, they are not yet identified, but existing (for example the Regional Portals); in part, finally, they will be realized in concert with the Portal. It is up to individual providers to define the concept of 'resource', choosing the information that is better representative of the entities that they want to make known through the Portal and protecting, if necessary, sensitive or confidential information. Therefore, an essential feature of the Portal metadata schema is that it supports the requirement of scalability, namely that the progressive introduction of new types of informational entities does not affect the operation of the entire system.

The Portal Application Profile

For compatibility with the choices made in the international arena, and in accordance with the guidelines established by MINERVA (the Ministerial Network for Valorising Activities in digitization, eContentplus - Supporting the European Digital Library), the import of metadata (i.e. information describing resources) from external data sources in the portal will take place with the mapping in a single scheme. This scheme is expressed according to the Qualified Dublin Core standard, which has been further extended to define the Application Profile designed for the Portal. The Qualified Dublin Core is a derived version of the Simple Dublin Core, encoded by the Dublin Core Metadata Initiative. It includes a set of basic elements with which any kind of resource can be described. A resource is 'anything that has identity'.

117 *http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/thesaurus/4.3/thesaurus_4.3.0.skos.xml*

118 *http://www.w3.org/2004/02/skos/*

119 http://www.culturaitalia.it/opencms/export/sites/culturaitalia/ attachments/documenti/picoap/picoap1.0.xml

documents can be accessed. In particular, a first passage was necessary to translate the ICCD documentation forms to PICO application profile (2007). Then, the PICO application profile was mapped to the more recent international data models such as the Europeana Data Model and the CIDOC conceptual reference model, described in Section 7.8.1.

6.3 Standardisation in the field of digital geographic information

The UN-GGIM document "A Guide to the Role of Standards in Geospatial Information Management", prepared cooperatively by the OGC (Open Geospatial Consortium), the ISO/TC 211 and the International Hydrographic Organization (IHO), introduces a maturity model in the development of SDI in three tiers. The first tier addresses the capability of sharing Maps over the Web. This implies the definition of standard services to publish maps over the Web (Web Mapping Services – WMS/ISO19128) and to enable their discovery (Catalogue service for the Web – CSW). The second tier refers the capability of sharing, integrating and using geographic information from different and multiple sources. This tier intends the support to the discovering, access and use of geographic information by the community for their own purpose. The third tier concentrates its activities in the availability and access of framework data and applications for their use on multiple platforms (desktop, mobile, or others). Finally, the last tier addresses the geographic knowledge ecosystem. It brings SDI into the Web of data. In this context, the Web can be compared to a Global geographic information database from which data from different sources are connected together providing additional value to the existing data. In this context, geographic information is moving from a data and information paradigm to a knowledge environment.[120]

The problem of harmonisation in the field of geographic knowledge is well known. Different aspects of the data must be considered in order to obtain a full interoperability between the maps. The reference and projection systems must be similar or it must be possible to convert the data from one to another. The source of represented data should be similar, in order to obtain a similar definition, accuracy, and shape of similar entities. The semantics of the data should be similar. These points very often differ even locally from one geodatabase to the neighbouring one. The difference increases when considering the boundary areas of two states, because the language in which the maps are built is usually different.

This is obviously a problem from different viewpoints: disciplines that should operate at international levels, such as environment science, require homogeneous maps as a base (for this reason for example the European Directive INSPIRE, described in Section 6.3.6, was born). Moreover, the reasoning and sharing potential of information technologies demand common semantics for their best realization (OGC standards move towards these aims).

Some more significant and basic standards are briefly described in the following subsections.

[120] The italic text cites (ISO, 2015)

6.3.1 ISO TC211 Geographic information / Geomatics

International Community – from 1994

Many bodies are actively engaged in the work of ISO/TC 211[121]. These include national standardization bodies, the OpenGIS Consortium (OGC), international professional bodies (such as FIG and ICA), UN agencies, and sectoral bodies (such as DGIWG and ICAO) (Figure 72).

The scope of the ISO technical commission 211 is the standardization in the field of digital geographic information as a support for decision-making where location (relative to the Earth) is a component.

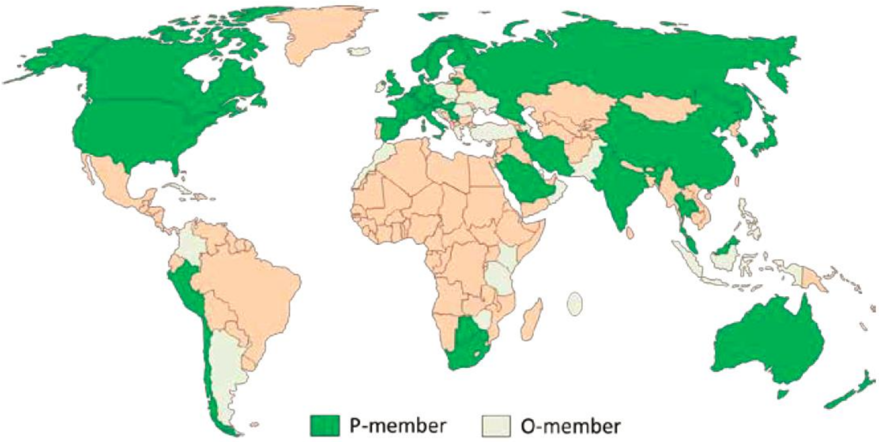


Figure 72. Participation in ISO/TC 211: ISO/TC 211 has participation from all continents – good participation from Europe, Asia and North America.

The work shall link to appropriate standards for information technology and data where possible, and provide a framework for the development of sector-specific applications using geographic data.[122]

The ultimate benefits of standardization are based on the use of widely recognized and accepted international voluntary standards developed to the highest technical level by an open consensus process that includes all those affected. These benefits occur at three levels.

1) *Those relating to the general benefit of standardization (particularly in areas of Information Systems and Technology):*
   a. reduced development costs of systems and applications, through established procedures;
   b. improved reliability of systems through reuse of mature components;
   c. reduced lock-in to individual proprietary suppliers.
2) *Those relating to standardization of geographic information at the general level:*
   a. ability to interoperate between systems through use of common schemas and interfaces;
   b. increased understanding of data (common terminology and standardised data definitions);
   c. ability to integrate locational data through use of standardised spatial referencing systems;
   d. ability to discover data resources through standardized metadata;
   e. improved data quality, through use of standardised quality measures;

[121] http://www.isotc211.org/
[122] The italic text cites http://www.isotc211.org/

    f. formalising standards produced by industry bodies, such as the Open Geospatial Consortium (OGC), for example GML (Geographical Markup Language);
    g. provision of models in a central repository to assist systems and applications developers.

3) T*hose relating to specific areas dealt with by TC 211:*
    a. location based services;
    b. observation and measurement;
    c. imagery sensor technology;
    d. land information systems.[123]

These standards may specify, for geographic information, methods, tools and services for data management (including definition and description), acquiring, processing, analysing, accessing, presenting and transferring such data in digital/electronic form between different users, systems and locations.

The scope of ISO/TC 211 is very wide, but the work currently concentrates in a few segments:

- *the modelling and documentation of geographic information;*
- *the spatial data infrastructures (SDIs) with emphasis on sharing and dissemination of geographic information through services;*
- *the embedding of geographic information in everyday life – ubiquitous geographic information;*
- *some specific domains where geographic information is an important component, and where multiple disciplines are involved.*

The main view of ISO/TC 211 is geographic information and geomatics as a horizontal and enabling technology – as an infrastructure – much like the focus of Information and communications technology (ICT).

In particular, this research should share the last scope, treating simple spatial information with the same geographic one with some similar approach for enhancing the representation.

A large number of projects are being realised for the publishing of standards for the different components of geomatics information.

The more interesting ones for the subject of this thesis are:

19101:2002 - Reference model (under revision) 19101-1 Reference model - Part 1: Fundamentals (revision)

19103 - Conceptual schema language (revision)

19104:2008 - Terminology

19110:2005 - Methodology for feature cataloguing (under revision)

19110:2005/Amd 1:2011z

19110 - Methodology for feature cataloguing (revision)19107:2003 - Spatial schema

19108:2002 - Temporal schema 19108/Cor 1

19136:2007 - Geography Markup Language

19137:2007 - Core profile of the spatial schema

[123] The italic text cites (ISO, 2015)

19144-1:2009 - Classification Systems – Part 1: Classification system structure

19142:2010 - Web Feature Service

19146:2010 - Cross-domain vocabulary

19150-1 - Ontology - Part 1: Framework: ISO/TS 19150-1:2012 defines the framework for semantic interoperability of geographic information. This framework defines a high level model of the components required to handle semantics in the ISO geographic information standards with the use of ontologies.

19150-2 - Ontology - Part 2: Rules for developing ontologies in the Web Ontology Language (OWL): ISO 19150-2:2015 defines rules and guidelines for the development of ontologies to support better the interoperability of geographic information over the Semantic Web. OWL is the language adopted for ontologies. It defines the conversion of the UML static view modelling elements used in the ISO geographic information standards into OWL. It further defines conversion rules for describing application schemas based on the General Feature Model defined in ISO 19109 into OWL. It does not define semantics operators, rules for service ontologies, and does not develop any ontology.

19125-1:2004 - Simple feature access - Part 1: Common architecture

19125-2:2004 - Simple feature access - Part 2: SQL option

19155 - Place Identifier (PI) Architecture

In particular, it's meaningful to underline some of the areas on which the ISO/TC is now working, since they are the same considered in this thesis, or concur in the definition of the application field.

***Ontologies**. A review summary report (N2705) recommended the development of a suite of standards in support of the Semantic Web. ISO/TS 19150-1, Geographic information - Ontology - Part 1: Framework, was published in 2012, and Part 2: Rules for developing ontologies in the Web Ontology Language (OWL) was published in 2015. Further standards on service ontology, content ontologies and semantic operators will follow.*

***Imagery encoding.** A Technical Specification on content components and encoding rules for imagery and gridded data is currently under development (ISO 19163), scheduled for publication in 2015-11. Interest in developing Technical Specifications for specific bindings to these content components and encoding rules have been expressed.*

A number of maintenance groups in ISO/TC 211 coordinate harmonization among ISO/TC 211 standards, as well as with the geographc information community at large:

- *The **Terminology Maintenance Group (TMG)** maintains the ISO/TC 211 terminology records and administers, coordinates, maintains and publishes a multi-lingual register of the terminology.*
- *The **Harmonized Model Maintenance Group (HMMG)** maintains the harmonized model to ensure that the UML models of ISO/TC 211 projects and standards are harmonized.*
- *The **XML Maintenance Group (XMG)** ensures that ISO/TC 211 XML is maintained and made accessible. The ISO/TC 211 XML represents any XML documents or fragments under the responsibility of ISO/TC 211, including the XML in published standards and specifications; in draft standards and specifications submitted for ballot; and XML resulting from the application of these (draft) standards and specifications.*
- *The **Group on Ontology Management (GOM)** ensures the development, maintenance and accessibility on the Web of ISO/TC 211 ontologies. This is done by setting a well-defined system of dereferenceable HTTP URIs for ISO/TC 211 concepts and their representations; by developing good practices for the use of Web Ontology Language (OWL) to describe data*

> *values of geographic information; by ensuring the development, maintenance and accessibility of OWL representations of geographic information concepts.*[124]

ISO/TC 211, through its Advisory Group on Outreach, seeks to promote the awareness, adoption, and advocacy of ISO/TC 211 standards in user communities.

6.3.2 Open Geospatial Consortium (OGC)

International Community – from 1994

The Open Geospatial Consortium (OGC) is a voluntary international not-for-profit organization committed to making quality open standards for the global geospatial community. These standards are made through a consensus process and are freely available for anyone to use to improve sharing of the world's geospatial data. Its members (more than 514) come from government, commercial organizations, NGOs (non-governmental organizations), *academic and research organizations.*[125]

The focus of OGC work is to define, document and test implementation standards for use with geospatial content and services. OGC standards leverage the abstract standards defined by ISO/TC 211. OGC develops standards that can be directly implemented. Many of these implementation standards are based on the conceptual (or abstract) models defined by ISO or jointly by the OGC and ISO. Therefore, the OGC maintains a Category A Liaison relationship with ISO/TC 211.[126]

The OGC has 4 core programs that enable to innovate and deliver high quality standards for the geospatial community.

1. ***Standards Program****: the Technical Committee and Planning Committee work in a formal consensus process to arrive at approved (or "adopted") OGC® standards.*
2. ***Interoperability Program****: global, innovative, hands-on prototyping and testing program designed to accelerate interface development and validation, and bring interoperability to the market.*
3. ***Compliance Program:*** *provides the resources, procedures, and policies for improving software implementations' compliance with OGC standards. The Compliance Program provides an online free testing facility.*
4. ***Communications and Outreach Program****: the OGC and its members offer resources to help technology developers and users take advantage of the OGC's open standards. Technical documents, training materials, test suites, reference implementations and other interoperability resources developed in OGC Interoperability Initiatives are available on the resources page http://wwwdev.opengeospatial.org/resources. In addition, the OGC and its members support publications, workshops, seminars and conferences to help technology developers, integrators and procurement managers introduce OGC plug and play capabilities into their architectures.*

Implementation Standards *are written for a more technical audience and detail the interface structure between software components. An interface specification is considered to be at the implementation level of detail if, when implemented by two different software engineers in ignorance of each other, the resulting components plug and play with each other at that interface. Any Schemas (xsd, xslt, etc)*

[124] The italic text cites (ISO, 2015)
[125] http://www.opengeospatial.org/
[126] The italic text cites (ISO, 2015)

that support an approved Implementation Standard can be found in the official OGC Schema Repository.

In 0, a selection of OGC implementation standards is reported, which is useful for the presented research.

The OGC Technical Committee (TC) has developed an architecture in support of its vision of geospatial technology and data interoperability called the OGC ***Abstract Specification****. The Abstract Specification provides the conceptual foundation for most OGC specification development activities. Open interfaces and protocols are built and referenced against the Abstract Specification, thus enabling interoperability between different brands and different kinds of spatial processing systems. The Abstract Specification provides a reference model for the development of OGC Implementation Standards.*

A selection of them are reported in Annex B.

For OGC (http://www.opengeospatial.org/resource/products/stats) there are currently 795 Implementing Products and 209 Compliant Products. Some of these are also ISO standards, many of the others are based on ISO 19107 and/or ISO 19125-1.

- ***Implementing Product****: a developer has obtained a copy of an OGC Standard and has made an attempt to implement it in software, follow the standard's instructions regarding interface or schema syntax and behaviours.*
- ***Compliant Product****: the OGC Compliance Testing Program provides a formal process for testing compliance of products that implement OGC Standards. A Compliance Testing determines that a specific product implementation of a particular OGC Standard complies with all mandatory elements as specified in the standard and that these elements operate as described in the standard.*[127]

6.3.3 OpenGIS® Geography Markup Language Encoding Standard (GML)

International Community – from 2007 ISO: 19136

The Geography Markup Language (GML) is an XML grammar for expressing geographical features. GML serves as a modelling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. As with most XML based grammars, there are two parts to the grammar:

- *the schema that describes the document;*
- *the instance document that contains the actual data.*

A GML document is described using a GML Schema. This allows users and developers to describe generic geographic data sets that contain points, lines and polygons. However, the developers of GML envision communities working to define community-specific application schemas that are specialized extensions of GML. If everyone in a community agrees to use the same schemas they can exchange data easily and be sure that a road is still a road when they view it. GML is also an ISO standard (ISO 19136:2007).[128]

[127] (ISO, 2015)

[128] The italic text cites http://www.opengeospatial.org/standards/gml

GML is the basic exchange format of geographic data. Being an extension of XML, the advantages and the mechanisms of this markup language can be exploited.

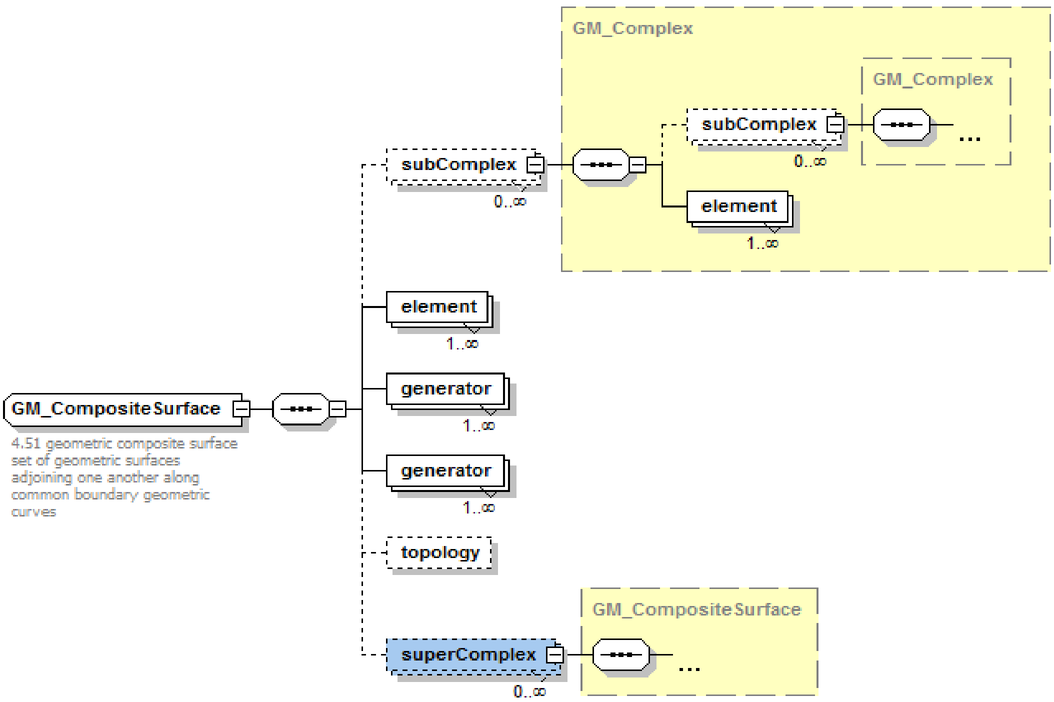
It is implemented and used for internet-based applications and data models. Another language for sharing geographic information on the web as linked open data was developed by OGC, it is named geoSPARQL. It is described in Section 7.9.1, and it refers to GML.

Its definitions already include the management of the topology and of the time. These are shared by ISO TC211 and INSPIRE (Annex C). Some meaningful UML schemas follow, that are defined in UML parts of the GML standard.

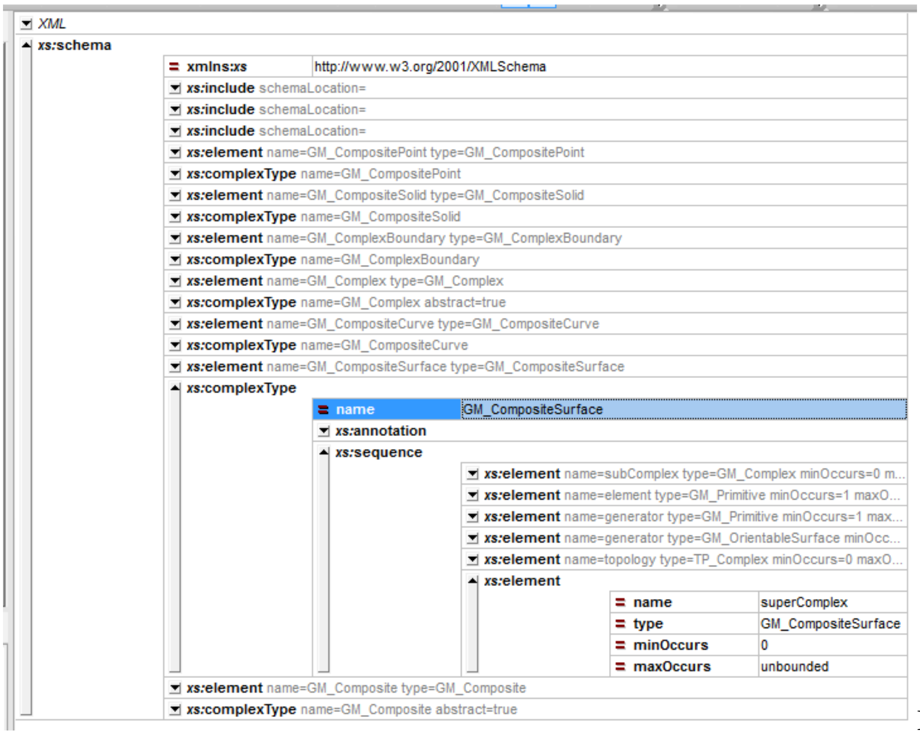
In the XML schema they appear as in Figure 73.

```xml
<xs:element name="GM_CompositeSurface" type="GM_CompositeSurface"/>
<xs:complexType name="GM_CompositeSurface">
  <xs:annotation>
    <xs:documentation>4.51 geometric composite surface
set of geometric surfaces adjoining one another along common boundary geometric curves</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="subComplex" type="GM_Complex" minOccurs="0" maxOccurs="unbounded"/>
    <xs:element name="element" type="GM_Primitive" minOccurs="1" maxOccurs="unbounded"/>
    <xs:element name="generator" type="GM_Primitive" minOccurs="1" maxOccurs="unbounded"/>
    <xs:element name="generator" type="GM_OrientableSurface" minOccurs="1" maxOccurs="unbounded"/>
    <xs:element name="topology" type="TP_Complex" minOccurs="0" maxOccurs="1"/>
    <xs:element name="superComplex" type="GM_CompositeSurface" minOccurs="0" maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>
```

I



II



III

Figure 73. (I) Example of XML hierarchy for geometric complex (in the XSD). (II) Tree – schema. (III) 'nested tables'.

6.3.4 OGC CityGML

International Community – ISO standard from 2008

CityGML is an open data model and XML-based format for the storage and exchange of virtual 3D city models. It is an application schema for the Geography Markup Language version 3.1.1 (GML3), the extendible international standard for spatial data exchange issued by the Open Geospatial Consortium (OGC) and the ISO TC211. The aim of the development of CityGML is to reach a common definition of the basic entities, attributes, and relations of a 3D city model. This is especially important with respect to the cost-effective sustainable maintenance of 3D city models, allowing the reuse of the same data in different application fields.

*The OGC Members adopted version 1.0.0 of CityGML as an official OGC Standard in August 2008. In late 2011, the OGC Members approved version CityGML 2.0.0 (*Figure 74*).*[129]

CityGML[130] is an open data model, an application-independent information model (generally provided in the form of application schema XSD) and an exchange format for GML files aimed at the representation, storage, and exchange of 3D urban objects, including city and landscape models. The original aim (its history begins in 2007) was to foster the reusability of 3D city models. Its semantic definition can be equally useful to manage the semantics of the data with the tools offered by informatics and artificial intelligence.

[129] The italic text cites http://www.opengeospatial.org/standards/citygml

[130] http://www.opengeospatial.org/standards/citygml

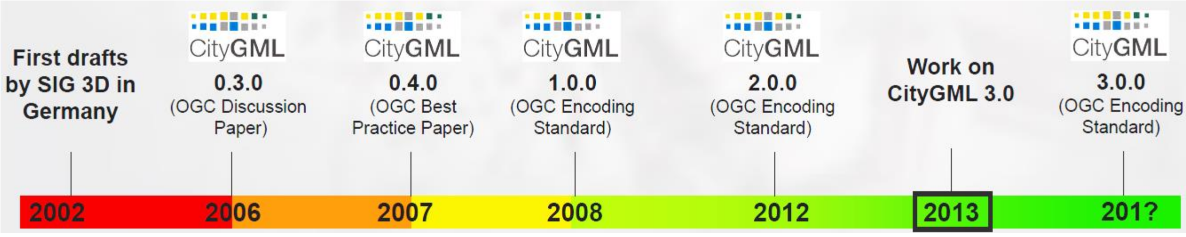


Figure 74. Timeline of CityGML development, from C. Nagel: CityGML – Urban Information Modelling, 21.01.2014). CityGML 3.0 should be approved in 2016[131].

CityGML is the most affirmed standard for building 3D models as a specification of urban cartography. It extends GML, specifying the semantic values of the records in GML for the 3D representation of city objects (Kolbe et al., 2005). The geometry models are therefore a subset of GML (Figure 134).

Each geometry is part of the entity defining the semantics, as is visible, for example, in the Building module in Figure 77. Several geometries can be associated to the same object for obtaining a multi-representation, based on time, on different reconstruction hypotheses, or different levels of detail. This last case is foreseen by the standard that implements 5 levels of detail (LoD) for a multi-scale representation of cartographic objects.

The concept of level of detail in the standard considers the accuracy of the represented features, which is indicative of the scale. As visible in the LoD 4 characteristics reported in Table 3, it generally respects a 0.2 m accuracy, which is used for 1:1000 representation scales. Nevertheless, the detail in the represented objects should be considered. This reaches a maximum definition as is typical of cartographic representations: a reference scale of 1:500 can be considered as maximum foreseen detail. The 1:500 representations can include objects that have a minimum approximately 10 cm dimension (as is approximately visible in the example in Figure 75). The scale 1:500 is also used for the maps of historical city centres, for which a higher detail was necessary, while remaining cartographic representations.

On the other hand, LoD 4 envisages the representation of interiors of buildings (including some furniture, see Figure 75). This is not usual in cartographic representation, where the maximum definition of some interior could be, on some occasions and for finite parts of the city, such as historical centres, the ground floor plants' schematic representations.

The application of a texture introduces other aspects for the characterization of the level of detail. In the texture, elements can be represented that have a larger detail (objects smaller than 10 cm) than the geometry. Therefore, a CityGML model in LoD 4 could carry more information relative to a traditional 1:1000 or 1:500 cartographic representation.

A first level of detail, LoD 0, deals with the representation of 2D or 2.5D (floating the 2D entities on a DTM or having the elevation value associated) entities. This one is generally intended for representing wide regions of land (scale about 1:50000).

The following LoDs have increasing-detail characteristics summarized in Table 3. An example of the LoDs is given in Figure 75.

These are insufficient for the representation of architectural heritage, since very often 1:200 and higher scales are needed, and sometimes it is useful to include 1:20 – 1:10 details, as well.

[131] https://github.com/opengeospatial/CityGML-3.0/wiki/CityGML3.0%20Overview

Table 3 Characterisitics of LoDs 1-4 representations in CityGML – Buildings[132]

	LoD 1	LoD 2	LoD 3	LoD 4
Model scale description	City, region	City district	Architectural models (outside), landmark	Architectural model (interior)
Class of accuracy	Low	Middle	High	Very high
Accuracy of position and height	5 m	2 m	0.5 m	0.2 m
Approximate representation scale	1:25000 – 1:10000	1:10000 – 1:5000	1:2500 – 1:1000	1:1000-1:500
Generalization	Object blocks as generalized features >6x6 m	Objects as generalized features >4x4 m	Object as real features >2x2 m	Constructive elements and openings are represented
Building installation			Representative exterior effects	Real object form
Roof form / structure	flat	Roof type and orientation	Real object form	Real object form

The inadequacy of CityGML model levels of detail for architectural heritage representation is evident if we compare them with the requirements for the recording levels of details stated by the CIPA document (Stylianidis et al., 2011). In that document, the levels of detail are classified following three main functions. The reconnaissance level of detail is not metric or has a small scale for understanding the object's overall general characteristics for a first identification of features and problems. The following level is the preliminary record, which must represent the resource's major features for a preliminary analysis, and it is required to have an accuracy of approximately ±10 cm for plans, elevations, and cross sections, and ±2 cm for structures and other structural details. This means that the suggested representation scale is approximately 1:50 or larger. This is already beyond what was foreseen by common CityGML applications, of which LoD4 is the largest one. In the CIPA document a further 'detailed record' is present, which must be the most accurate and complete record useful for the preservation aims and needs. For this one an accuracy of ±5 to 25 mm is recommended.

The standard provides data models of city objects (structured as shown by the UML model in Figure 76). One of the advantages of the model is that the modules can be extended or added for adapting to the different application fields needs.

[132] (Fan, Meng, 2009)

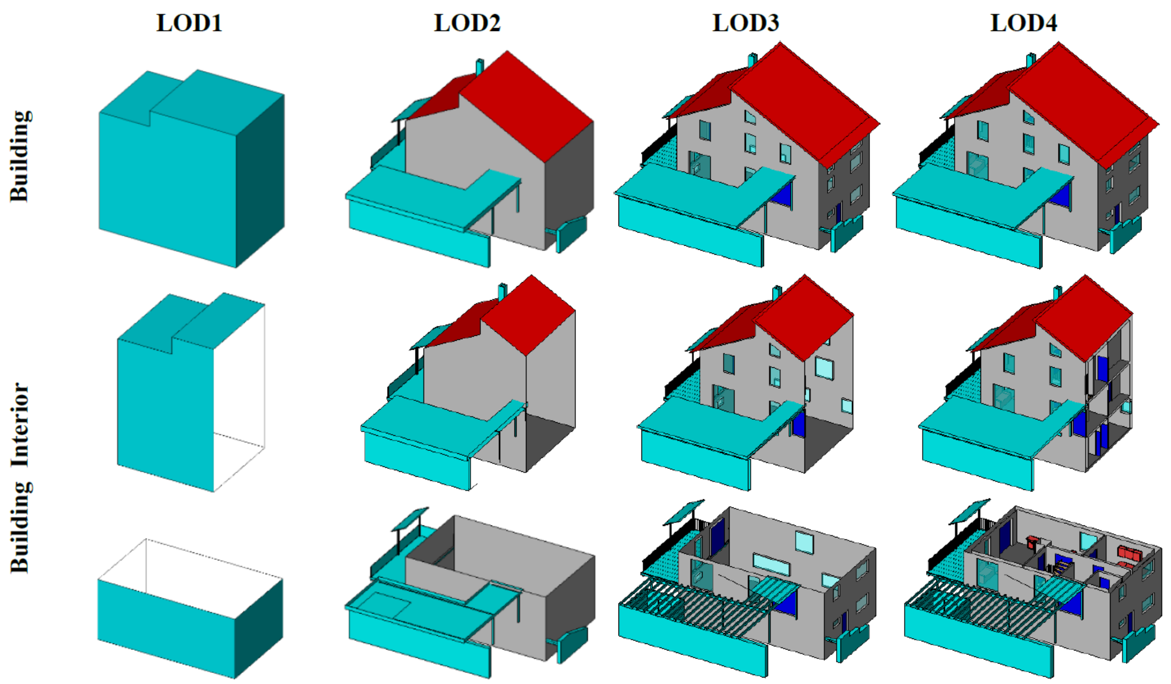


Figure 75. Building model in LoD1 – LoD4 (OGC, 2012)

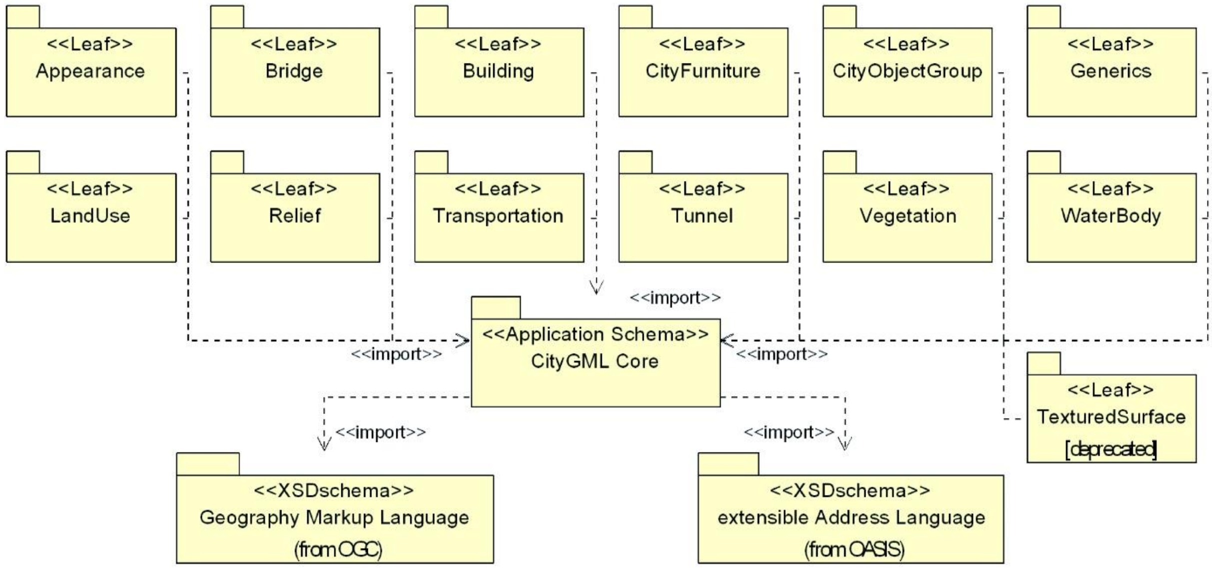


Figure 76. CityGML modules overview (OGC, 2012)

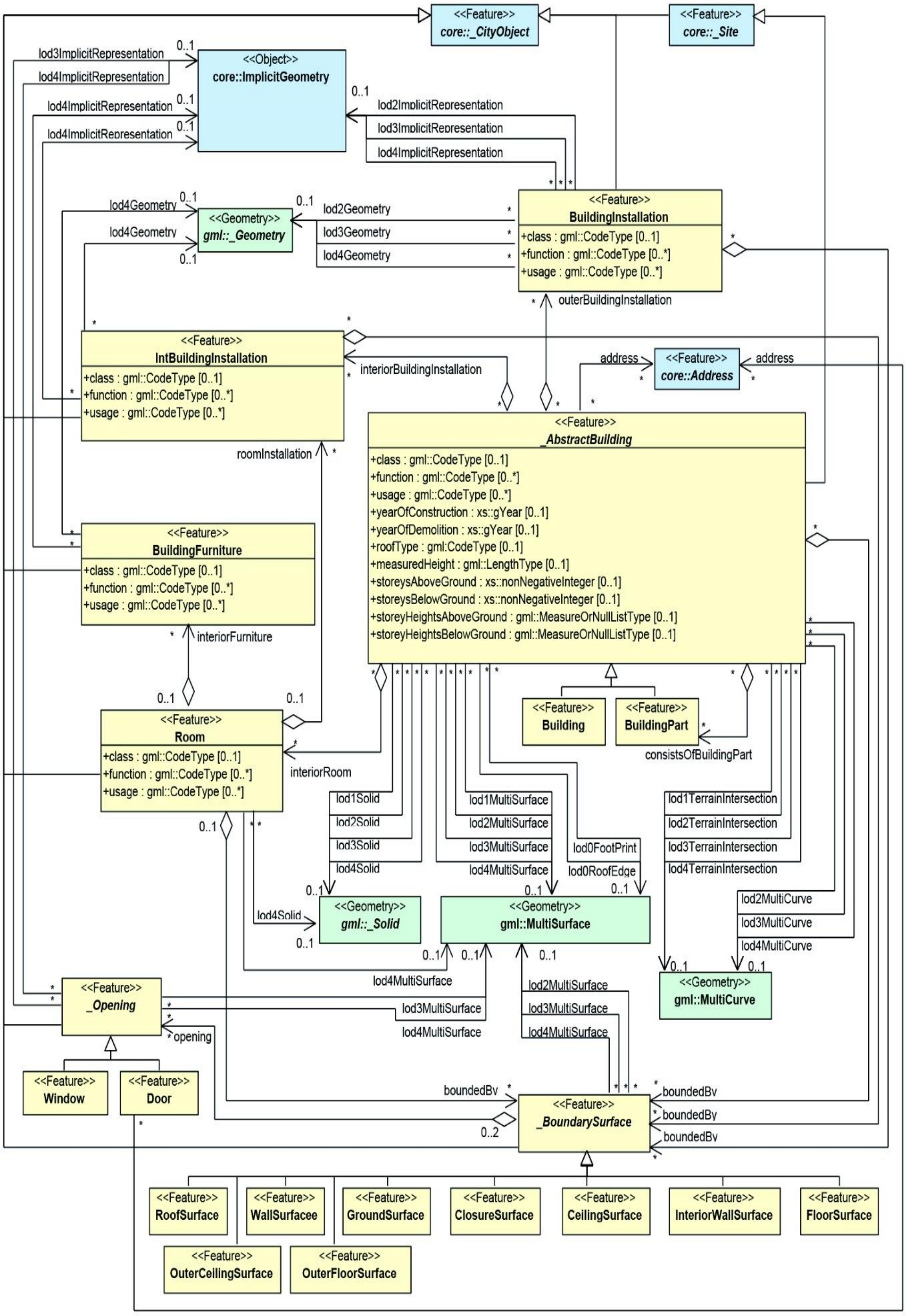


Figure 77. UML schema of the CityGML Building module (OGC, 2012).

6.3.5 CityGML Application Domain Extensions (ADE)

The defined CityGML model can be extended in order to model further aspects linked to specific application domains. The so-composed extensions use specific characteristics and procedures of CityGML, being defined as Application Domain Extension (ADE).

Some official ADEs exist[133]. They relate especially to some urban-scale themes, such as noise or inclusive routing. Some of these are specific to buildings; for example, GeoBIM integrates some classes derived from the IFC (Industry Foundation Classes)[134] standard used in BIM (Building Information Modelling) (de Laat, van Berlo, 2011). However, even if in the future the field of BIM is likely to meet GML models, at present it is too rigid for describing cultural heritage buildings, which require more flexibility.

A further investigation has been performed for the extension of the CityGML model in order to include some information about the cultural heritage (CH) nature of the building and some surface characteristics, such as the deterioration (Costamagna, Spanò, 2013).

The structures that have to be supportive of this mechanism are provided in the same CityGML specifications, and are reported in the following part. The possibility of using the ADEs empowers this standard by opening its representation possibilities with the eventuality of integration for adapting to further application fields.

In general, there are two different approaches to combine city model data and application data:

1) embed the CityGML objects into a (larger) application framework and establish the connection between application data and CityGML data within the application framework. For example, CityGML data fragments may be embedded into the application's XML data files or stored as attributes of application objects according to the application data model;
2) incorporate application-specific information into the CityGML instance documents. This approach is especially feasible, if the application-specific information follows essentially the same structure as defined by the CityGML schema. This is the case, if the application data could be represented by additional attributes of CityGML objects and only some new feature types would have to be defined.

The second approach is that one chosen for the application in the thesis research and also the one ruled by the CityGML specification. This one is therefore detailed.

The possibility to support extensions of the model is offered by the existence of some generic attributes and objects.

They can be used without changing the data model to add further attributes or objects to the archive. However, they have some disadvantages among which the worst is that they lack formal specification of the names, datatypes, and multiplicities, that is, their semantics is uncertain and interoperability suffers.

An additional formal XML schema based on the CityGML schema definitions (called CityGML *Application Domain Extension* - ADE) is desirable to represent application-specific information. It allows the validation of instance documents both against the extended CityGML and the ADE schema and therefore helps to maintain semantic and syntactic interoperability between different systems working in the same application field. In order to prevent naming conflicts, every ADE has to be

[133] http://www.citygmlwiki.org/index.php/CityGML-ADEs
[134] https://en.wikipedia.org/wiki/Industry_Foundation_Classes

defined within its own namespace, which must differ from the namespaces associated with the CityGML modules. An ADE schema may extend one or more CityGML module schemas. The relevant CityGML module schemas have to be imported by the ADE. Technical details on how an ADE can be defined are described in Section 6.3.5.1.

6.3.5.1 Technical principle of ADEs

Each ADE is specified by its own XML schema file. The target namespace is provided by the information community that specifies the CityGML ADE. The target namespace is a URL, i.e. a sort of link with unique identifier, to the path where the schema is stored (see Section 7.3 for URL and URI definitions). This is typically not the OGC or the SIG 3D (Spatial Interest Group for 3D city models development[135]). The namespace should be in the control of this information community and must be given as a previously unused and globally unique URI. This URI will be used in CityGML ADE instance documents to distinguish extensions from CityGML base elements. As the URI refers to the information community, it also denotes the originator of the employed ADE.

The XML file of an ADE schema must be available (or accessible on the Internet) to everybody creating and parsing CityGML instance documents, including these ADE-specific augmentations.

An ADE XML schema can define various extensions to CityGML. However, all extensions shall belong to one of the two following categories:

1. new feature types defined within the ADE namespace and based on CityGML classes. In general new feature types have to be derived from existing (here, CityGML) feature types. For example, new feature types could be defined by classes derived from the classes like *_CityObject* or *_AbstractBuilding*. The new feature types then automatically inherit all properties (i.e. attributes) and associations (i.e. relations) from the respective CityGML superclasses;

2. existing CityGML feature types are extended by application of specific properties (in the ADE namespace). These properties may have simple or complex data types. Also, geometries or embedded features (feature properties) are possible. The latter can also be used to model relations to other features.

In this case, the extension of the CityGML feature type is not being realised by the inheritance mechanism of the XML schema. Instead, every CityGML feature type provides a ‘hook’ in its XML schema definition that allows the attachment of additional properties to it by ADEs. This ‘hook’ is implemented as a GML property of the form ‘_GenericApplicationPropertyOf<*Featuretypename*>’ where <*Featuretypename*> is equal to the name of the feature type definition in which it is included. The datatype for these kinds of properties is always ‘xsd:anyType’ from the XSD namespace. The minimum occurrence of the ‘_GenericApplicationPropertyOf<*Featuretypename*>’ is 0, and the maximum occurrence is unbounded. This means that the CityGML schema allows that every CityGML feature may have an arbitrary number of additional properties with arbitrary XML content with the name ‘_GenericApplication-PropertyOf<*Featuretypename*>’. For example, the last property in the definition of the CityGML feature type *LandUse* is the element ‘*_GenericApplicationPropertyOfLandUse’*.

Such properties are called ‘hooks’ to attach application specific properties, because they are used as the head of a substitution group by ADEs. Whenever an ADE wants to add an extra property to an existing CityGML feature type, it should declare the respective element with the appropriate datatype

[135] http://www.sig3d.org/index.php?&language=en

within the ADE namespace. In the element declaration, this element shall be explicitly assigned to the substitution group defined by the corresponding '_GenericApplicationPropertyOf<*Featuretypename*>' in the corresponding CityGML module namespace. An example is given in the following subsection.

By following this concept, it is possible to specify different ADEs for different information communities. Every ADE may add its specific properties to the same CityGML feature type as they all can belong to the same substitution group. This allows the possibility of having CityGML instance documents where CityGML features contain additional information from different ADEs simultaneously.

That usage of ADEs introduces an extra level of complexity as data files may contain mixed information (features, properties) from different namespaces, not only from the GML and CityGML module namespaces. However, extended CityGML instance documents are quite easy to handle by applications that are not 'schema-aware', i.e. applications that do not parse and interpret GML application schemas in a generic way. These applications can simply skip anything from a CityGML instance document that is not from a CityGML module or GML namespace. Thus, a building is still represented by the *<bldg:Building>* element with the standard CityGML properties, but with possibly some extra properties from different namespaces. Also, features from a different namespace than those declared by CityGML modules or GML could be skipped (e.g. by a viewer application).[136]

The solution to this for the viewer application used in this thesis case study was to add the new schemas in the software files themselves. However, it should be noted that this is not always possible.

6.3.5.1.1 The OGC best practice for ADE construction: from UML to XSD

The affirmed best practice to build ADEs is the use of existing CityGML UML diagrams as bases in specific software and the processing of their extension with new classes in UML format (van den Brink et al., 2012). The application of stereotypes permits the definition of the nature of the destination XSD (XML Scheme Definition, that is, the scheme that structures the data in XML files) element and enables the automatic generation of the application schema.

This can be done through already implemented algorithms in the available software such as '*Enterprise Architect'* or '*ShapeChange'* applications (the second one is an open source software but it requires a series of further tools such as *TortoiseSVN, SVN checkout,* and so on).

The procedure is nevertheless affirmed; it is used for the compiling of standards (INSPIRE, CityGML) and is described by a specific annex (Annex E) of the GML encoding standards (OGC, 2007). It is also the one recommended in the INSPIRE Directive.

In particular, following the OGC best practice for extending an existing class with further attributes, a subclass having the same name of the class to be extended and stereotype 'ADEElement' should be created. The specialisation relation is marked with stereotype 'ADE'. For adding a new class, a simple subclass having stereotype 'featureType' must be added.

[136] The text cites (OGC, 2012)

6.3.6 INSPIRE European Directive

INSPIRE (Infrastructure for Spatial Information in the European Community)[137] is a Directive of the European Parliament (Directive 2007/2/EC) and of the Council (14 March 2007) establishing an infrastructure for spatial information to support a unique Community underground for the sharing of environmental spatial information among European States. The aim is to obtain a harmonized spatial information as reference for the Community environmental policies and activities that may have an impact on the environment. It entered into force on the 15th May 2007 and is supposed to be fully implemented by 2019.

Two of the main principles of INSPIRE are:

"It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications;

[…] Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used".

The INSPIRE Directive gives some basic guidelines for the production and the updating of the maps or further cartographic products, such as the definition of the reference system to be used (ETRF2000 – UTM). Moreover, it also provides a Conceptual Model including 34 spatial data themes.

The level of detail is lower than, for example, in CityGML, since the object of interest of INSPIRE are wide areas with transboundary dimension and, consequently, as explained in the introduction, being that the objects of interest differ, the represented entities differ.

Nevertheless, some common entities exist (for example, the entity 'Building'). Therefore, some harmonization activity is being realized. All the INSPIRE regulations and technical specifications already implement ISO TC211 in their definitions.

[137] http://inspire.jrc.ec.europa.eu/

7 SMART COMMUNITIES AND THE WEB EVOLUTION FOR INFORMATION RETRIEVAL

The ‘smart’ concept, primarily applied to cities, emerge at the end of 1990s and was acknowledged from 2000, with the beginning of an urban development programme in New York State in association with universities (Bowerman et al., 2000). It has been until now widely discussed and has continued to develop in active communities. The goals of smart cities are various, from the green buildings to the improved mobility.

Nevertheless, the preliminary aim to strive for the realization of further ambitions is the re-thinking and improvement of the relationships among government, city managers, business, academia, and research communities (Bowerman et al., 2000). This list has then been integrated with citizens/users, given the importance of the crowd in the processes of decision-making, planning, and, obviously, functioning of the cities (in an extended concept). The concept thus acquires a three-dimensional value: institutions, people, and, finally, technology (Nam, Pardo, 2011). The cited ambitions therefore change in the following areas: integration of infrastructures and technology-mediated services, social learning for strengthening human infrastructure, governance for institutional improvement and citizen engagement, and a number of more specific or operational nuances which evolve in parallel to technology (Nam, Pardo, 2011).

The technology thus plays a fundamental role, until it is mistaken for the real aim. However, the use of information and communication technology to sense, analyse, and integrate the information and systems in cities (Su et al., 2011) is the tool to reach the reciprocal interconnection among data providers, services, urban events, urban facts, and citizens’ needs or interests.

An example of this technology, basic to the realization of a smart digital society, is the ‘Internet of Things’ (IoT) (Zanella et al., 2014). It is an emergent concept that refers to the networked interconnection of everyday objects, which, in a trivial definition, consists in having multiple sensors connected to the internet that monitor the physical world and interact with each other, making possible through services to access remotely the data and to control the physical world from a distance.

The effective exploiting of the Internet of Things and the other structures of a smart city requires the correct interpretation of the published data to be known by people and machines. Heterogeneous ‘things’ must be integrated in a common frame in order to be understood, managed, and retrieved. The definition of the meaning of the data for enabling interoperability and interchange of unambiguous knowledge is the main goal shared by smart city promoters, ontologies, and standards and the Semantic Web (presented in Subsection 7.1). The developments of web technologies (archiving tools, big data, etc.) and the artificial intelligence field (neural networks, machine learning, etc.) make the realization of these ideas possible today. That is essential for an effective information structure, which could make the smart things a reality, in association with the IoT (Barnaghi et al., 2012). A preliminary condition for the achievement of such an objective is the application of semantics to the managed data (Rowley, 2007). In the light of premises, ontologies can serve as effective tools to structure and develop smart systems. A large community is including them when working for the goals of smart cities (Al-Hader, 2009; Schneider et al., 2012).

Nevertheless, geoinformation is fundamental for smart data retrieval and analysis (Prandi et al. 2014). Structures for an improved thematic plus spatial and geographic knowledge are therefore encouraged for the creation and use of smart archives.

Cultural heritage is not a secondary issue of smart cities, when considering identity and social issues, and the civilization value is not to be lost in digital society but, instead, should be preserved, shared, and brought out.

Some additional investigations are being developed on these themes (e.g. Amato et al., 2012; Ryan et al., 2005).

7.1 The Semantic Web

The World Wide Web[138] allows humans to seek information through documents based on word correspondences. The capability of machines to access the meaning of Web content and to interpret sentences and information is still very limited. The most valuable tools of the World Wide Web are search engines. However, some problems are associated with their use: the recall is not always high or sufficiently precise; the results are highly sensitive to the vocabulary, so they do not consider the semantic value of different words that could have similar meaning and vice versa. For example, the search for ‘capital’ can give results about: financial capital, the topmost member of a column, the town seat of the government of a State, an Italian radio station, and so on. Moreover, the results are single web-pages, so you must recollect the data spread in different documents in order to compile complete information (Antoniou, van Harmelen, 2004).

For the Web content to become more easily machine-processable, the World Wide Web is evolving towards the Semantic Web[139]. In particular, Tim Berners-Lee, the very person who invented the WWW in the late 1980s, is the driving force of the Semantic Web initiative. He expects from this initiative the realization of his original vision of the Web, a vision where the meaning of information played a far more important role than it does in the present-day Web (Antoniou, van Harmelen, 2004).

The Semantic Web is the evolution of the World Wide Web that permits the management of the data themselves, instead of documents. The developments of web technologies and the artificial intelligence field make its realization possible today. A preliminary condition for the achievement of such an objective is the application of semantics to the managed data (Rowley, 2007). In fact, the effective exploitation of the enhanced possibilities offered by the web, including huge volumes of data, mobile devices for both input and output of the data, and so on (Domingue, 2011), requires the correct interpretation of the published data to be known by people and machines. This complexity also includes the ‘IoT’: heterogeneous ‘things’ must be integrated in a common frame in order to be understood, managed and retrieved. The definition of the meaning of the data for enabling interoperability and interchange of unambiguous knowledge is the main goal of the Semantic Web.

In addition, Web searches can exploit generalized/specialized information. If a query fails to find any relevant documents, the search engine may suggest to the user a more general query. It is even conceivable for the engine to run such queries proactively to reduce the response time in case the user adopts a suggestion. Or if too many answers are retrieved, the search engine may suggest to the user some specializations. The system interacts with the user, offering interpretations and suggesting research paths. In Artificial Intelligence (AI), there is a long tradition of developing and using ontological languages (Antoniou, van Harmelen, 2004) (Figure 78).

[138] http://webfoundation.org/

[139] http://www.w3.org/standards/semanticweb

The available tools to realize it are provided by the organizations involved in the Semantic Web development, especially W3C (World Wide Web Consortium)[140], which developed the Semantic Web technologies. Among these, essential tools are ontologies and standard data models, that specify in a clear, unambiguous, and 'open' way to structure the data and to prescribe how they have to be interpreted.

Semantic Web agents will make use of all the technologies we have outlined:

- *Metadata* will be used to identify and extract information from Web sources;
- *Ontologies* will be used to assist in Web searches, to interpret retrieved information, and to communicate with other agents. Also, web services should take advantage of them (Roman et al., 2005). Ontologies are useful for improving the accuracy of Web searches. The search engines can look for pages that refer to a precise concept in an ontology instead of collecting all pages in which certain, generally ambiguous, keywords occur. In this way, differences in terminology between Web pages and the queries can be overcome;
- *Vocabularies* define the concepts and relationships (also referred to as 'terms') used to describe and represent an area of concern. Vocabularies are used to classify the terms that can be used in a particular application, characterize possible relationships, and define possible constraints on using those terms. The role of vocabularies on the Semantic Web are to assist data integration when, for example, ambiguities may exist on the terms used in the different data sets, or when a bit of extra knowledge may lead to the discovery of new relationships;[141]
- *Logic* will be used for processing retrieved information and for drawing conclusions.

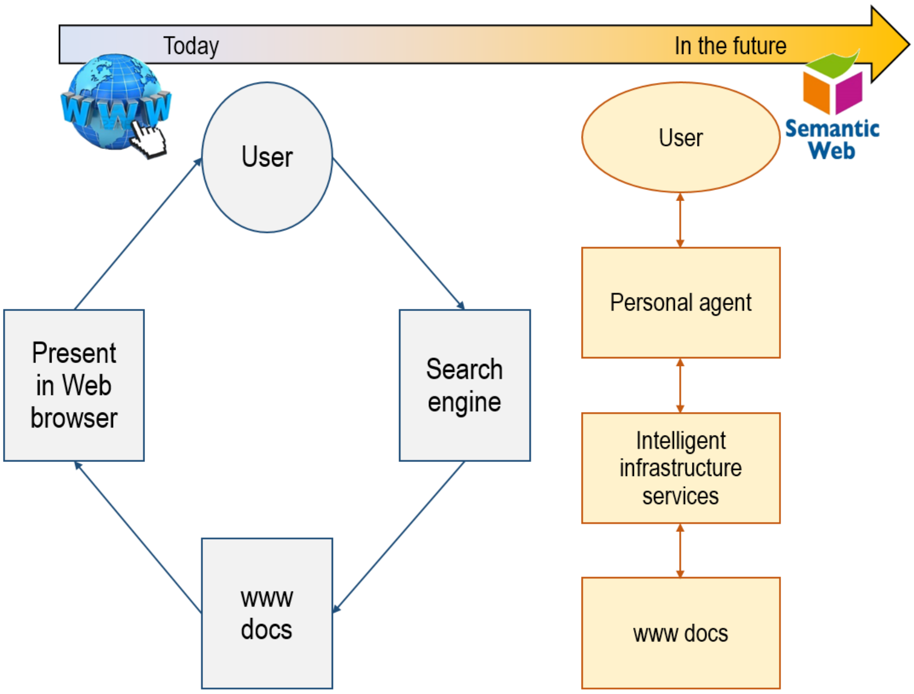


Figure 78. Mechanisms for data retrieval from the web in the World Wide Web (left) and in the Semantic Web (right).

In addition to the classic 'Web of documents', W3C is helping to build a technology stack to support a 'Web of data', the sort of data found in databases. The ultimate goal of the Web of data is to enable

[140] www.w3.org
[141] https://www.w3.org/standards/semanticweb/data

computers to do more useful work and to develop systems that can support trusted interactions over the network. The term ‘Semantic Web’ refers to W3C’s vision of the Web of linked data. Semantic Web technologies enable people to create data stores on the Web, build vocabularies, and write rules for handling data. Linked data are empowered by technologies such as RDF, SPARQL, OWL, and SKOS (Simple Knowledge Organization System[142]) (Figure 79).

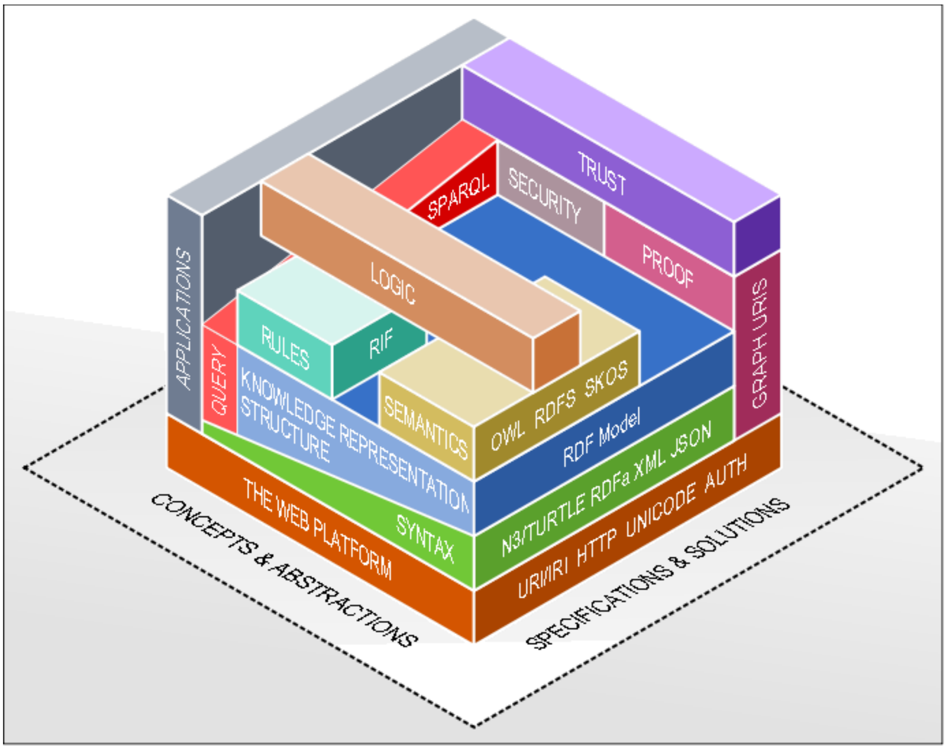


Figure 79. The layered Semantic Web technology stack[143].

An important issue of the Semantic Web is the inference: reasoning over data through rules. W3C works on rules, primarily through RIF and OWL, it is focused on translating between rule languages and exchanging rules among different systems. Broadly speaking, inference on the Semantic Web can be characterized by discovering new relationships. On the Semantic Web, data are modelled as a set of (named) relationships between resources. ‘Inference’ means that automatic procedures can generate new relationships based on the data and based on some additional information in the form of a vocabulary. Inference on the Semantic Web is one of the tools of choice to improve the quality of data integration on the Web, by discovering new relationships, automatically analysing the content of the data, or managing knowledge on the Web in general. Inference-based techniques are also important in discovering possible inconsistencies in the (integrated) data.[144]

In order to include architectural heritage documentation, including 3D models, in this system to take advantage of the great power of the studies, the data model must be considered and standard XML-based languages must be used.

[142] https://www.w3.org/2004/02/skos/
[143] https://smiy.wordpress.com/2011/01/10/the-common-layered-semantic-web-technology-stack/
[144] https://www.w3.org/standards/semanticweb/data

7.2 The Linked Open Data

The result of the realization of the Semantic Web are the so-called 'Linked Open Data' and, in turn, the Linked Open Data make the Semantic Web (the Web of Data) a reality.

It is important to have the huge amount of data on the Web available in a standard format, reachable and manageable by Semantic Web tools. Furthermore, the Semantic Web not only requires access to data, but *relationships among data* should be made available, as well, to create a *Web* of Data (as opposed to a sheer collection of datasets). This collection of interrelated datasets on the Web can also be referred to as Linked Data.

To achieve and create Linked Data, technologies should be available for a common format (RDF, see Section 7.5.1), to facilitate either conversion or on-the-fly access to existing databases (relational, XML, HTML, GML, etc.). It is also important to be able to set up query endpoints to access that data more conveniently (W3C provides technologies for this aim).

Linked Data lies at the heart of what the Semantic Web is all about: large-scale integration of, and reasoning on, data on the Web.[145]

Why Linked Open Data?

When data is linked and open, data is structured and published according to principles allowing it to be interlinked and made openly accessible and shareable on the Semantic Web

Linked Open Data (LOD) on the Web has the potential to connect data from different domains and can use the Web to create defined "typed" links for meaningful linking between data from various resources.

LOD will enable a new generation of search engines to follow the links between data resources in order to deliver ever more complete answers as new data sources become available.

In a LOD world, users' research will operate on top of an unbound, global data space.

LOD will facilitate the ability to derive new information from the information the user already has, through inference, classification and other applications of logic onto the data.

Publishing information as LOD unlocks the full knowledge potential of databases, allowing researchers to access, query, and recombine data created and stored by institutions of all kinds (libraries, museums, archives, research centres, laboratories, universities, governments, publishers, and others).[146]

Key values to Linked Open Data:

- *everyone can access, reuse, enrich and share the data published in LOD format. It makes "interworkability" more feasible;*
- *LOD disambiguates your search. In other words it eliminates the noise you sometimes get in web search engine search. What you are seeking is more precise because of the way you tag the information you are converting to LOD;*
- *LOD connects you to all kinds of relevant information creating a playground for serendipity and it updates that information dynamically within the same web space;*

[145] https://www.w3.org/standards/semanticweb/data

[146] The italic text cites http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf

- *LOD is able to connect data from the Web that was not previously connected.*[147]

7.3 The concept of Uniform Resource Identifier (URI)

In informatics, a Uniform Resource Identifier refers to a string uniquely identifying a resource (a document, a concept, a service, etc.). It is essential for building data connections, since it permits the unique identification of each resource and does not permit ambiguities in the data archiving and in the data retrieval.

It can be classified as:

- URL (Uniform Resource Locator) that permits to access to the resource;
- or URN (Uniform Resource Name) that permits the identification of a resource by a name (normally a domain name or 'namespace'); it is similar to the unique ID (identifier) in databases.

Some rules exist for evaluating their efficiency, and some guidelines exist for their formulation[148]. In particular, they consider the URI to be both machine- and human-readable: identifiers are not only used by computers, but also by humans. Humans have a strong preference for using mnemonic names. Such identifiers are easier to create, easier to remember, easier to understand, easier to guess, easier to transcribe, and easier to identify with.[149]

The URI permits the establishment of connections and relations among data and consequently make the realization of a web of data possible (it is used by all the semantic web tools, such as RDF, OWL, SPARQL, Xlink and so on).

A particular type of URI is the Internationalized Resource Identifiers (IRI): this is a protocol element, a complement to URIs. An IRI is a sequence of characters from the Universal Character Set (Unicode/ISO10646). There is a mapping from IRIs to URIs, which means that IRIs can be used instead of URIs, where appropriate, to identify resources.

Perhaps the most important characteristic of IRIs in web architecture is that they can be dereferenced, and hence serve as starting points for interactions with a remote server. This specification is not concerned with such interactions. It does not define an interaction model. It only treats IRIs as globally unique identifiers in a graph data model that describes resources. However, those interactions are critical to the concept of Linked Data that makes use of the RDF data model and serialization formats.[150]

[147] The italic text cites http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf
[148] https://www.w3.org/Addressing/URL/uri-spec.html
[149] https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
[150] https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/

7.4 W3C (World Wide Web Consortium)

The World Wide Web Consortium (W3C)[151] *is an international community where Member organizations, a full-time staff and the public work together to develop Web standards. Led by Web inventor Tim Berners-Lee and CEO Jeffrey Jaffe, W3C's mission is to lead the Web to its full potential.*[152]

The W3C mission is to lead the World Wide Web to its full potential by developing protocols and guidelines that ensure the long-term growth of the Web. Below we discuss important aspects of this mission, all of which further W3C's vision of One Web.

The following **design principles** guide W3C's work.

- *Web for All:*

The social value of the Web is that it enables human communication, commerce, and opportunities to share knowledge. One of W3C's primary goals is to make these benefits available to all people.

- *Web on Everything*

The number of different kinds of devices that can access the Web has grown (mobile phones, smart phones, personal digital assistants, interactive television systems etc.)

W3C's **vision** for the Web involves participation, sharing knowledge, and thereby building trust on a global scale.

- *Web for Rich Interaction:*

The Web was invented as a communications tool intended to allow anyone, anywhere to share information.

- *Web of Data and Services:*

Some people view the Web as a giant repository of linked data while others as a giant set of services that exchange messages. The two views are complementary, and which one to use often depends on the application.

- *Web of Trust:*

The Web transformed the way we communicate with each other. In doing so, it has also modified the nature of our social relationships. People now "meet on the Web" and carry out commercial and personal relationships, in some cases without ever meeting in person. W3C recognizes that trust is a social phenomenon, but technology design can foster trust and confidence. As more activity moves on-line, it will become even more important to support complex interactions among parties around the globe.

W3C publishes therefore the recommendations and standards which are crucial to the realization of the Semantic Web (XML, RDF, OWL, SPARQL, etc.) (7.4.1, 7.5).[153] W3C adhere to Open Stand (7.6).

[151] *https://www.w3.org/*
[152] The italic text cites https://www.w3.org/
[153] The italic text cites https://www.w3.org/

7.4.1 XML technologies

EXtensible Markup Language (XML)[154] is a W3C Recommendation[155], designed to store and transport data among applications. It derives from SGML (ISO 8879) and was designed for large-scale electronic publishing. It has the advantage to be both human- and machine-readable[156].

It can also be used to structure internet documents. It differs from HTML (HyperText Markup Language) since they were designed with different goals: XML was designed to carry data, with a focus on what data *is*, whereas HTML was designed to display data, with a focus on how data *looks*. Moreover, XML tags are not predefined like HTML tags are.

The tags can be defined by the author of the XML document. This is the reason why the interchange XML files produced by some proprietary software (e.g. *ESRI ArcGIS*) often cannot be read by different applications. However, this also improves the flexibility of the language, enabling the possibility to extend it.

It is one of the key enabling technologies of the SemanticWeb.

XML documents are structured by further specific XML documents, called 'XML schema definition' (Section 7.4.1.1), and exploit several mechanisms for operating on them by means of links, translations, and further interactions (Figure 80).

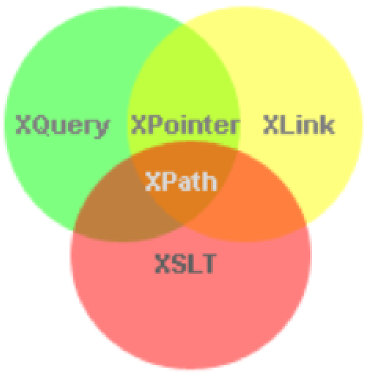


Figure 80. Mechanisms used to operate with XML technologies.

XML is the base of several languages used to structure and share data on the web and among software tools and applications. The more meaningful ones to this research are summarized in Table 4, and described in following chapters.

Table 4 Summary of the described XML-based languages and technologies.

<table>
<tr><td></td><td colspan="2">XML</td><td colspan="2">XML extensions</td></tr>
<tr><td rowspan="5">XML tools</td><td>XPath</td><td rowspan="5"></td><td>XSD</td><td rowspan="3">Schema definition</td></tr>
<tr><td>XQuery</td><td>RDFS (RDF extension)</td></tr>
<tr><td>XLink</td><td>OWL (RDF extension)</td></tr>
<tr><td>XPointer</td><td>GML</td><td rowspan="2">Structured data</td></tr>
<tr><td>XSLT</td><td>RDF</td></tr>
</table>

[154] https://www.w3.org/XML/
[155] https://www.w3.org/TR/2008/REC-xml-20081126/
[156] http://www.w3schools.com/xml/

7.4.1.1 XML Schema Definition (XSD)

XSD is a recommendation of the W3C (published in 2001) to formally describe the elements in an XML document. It uses, in turn, XML, and can be used to express a set of rules to which an XML document must conform in order to be considered 'valid' according to that schema.[157]

The data model definitions for structuring XML databases uses XSD encoding.

7.4.1.2 XPath

XPath (1999) is used to navigate through elements and attributes in an XML document. It is a syntax for defining parts of an XML document in order to navigate in them using path expressions.

*Moreover, it is a major element in XSLT (*EXtensible Stylesheet Language for Transformations)(7.4.1.6).[158]

7.4.1.3 XQuery

The W3C recommendations XQuery (2007) is the language for querying XML data. It is for XML like SQL is for databases. XQuery is a language for finding and extracting elements and attributes from XML documents.

XQuery is built on XPath expressions.[159]

The selections in XQuey can be done using the "FLWOR" expression instead of the "select – from – where" typical of SQL selections[160]

FLWOR is an acronym for "For, Let, Where, Order by, Return":

- the "for" clause selects all the elements into a variable called $x;
- the "where" clause selects the elements with a property element respecting some condition;
- the order by clause defines the sort-order;
- the return clause specifies what should be returned.

7.4.1.4 Xlink

XLink (1999) is used to create hyperlinks within XML documents.[161] It can be used for linking both external resources (by means of a URI) and internal elements (by means of ID).

In the thesis application, it can be effectively used for defining relations among objects (it is also used for establishing topology relationships) in the interior of one document. It can also be used to set the destination of external references to further resources (such as external vocabularies).

[157] https://en.wikipedia.org/wiki/ XML_Schema_(W3C)
[158] The italic text cites http://www.w3schools.com/xsl/xpath_intro.asp
[159]The italic text cites http://www.w3schools.com/xsl/xquery_intro.asp
[160] http://www.w3schools.com/xsl/xquery_flwor.asp
[161] http://www.w3schools.com/xml/ xml_xlink.asp

7.4.1.5 XPointer

XPointer (1999) allows links (in association with XLink) to point to specific parts of an XML document, by using XPath expressions[162]

7.4.1.6 EXtensible Stylesheet Language for Transformations (XSLT)

The W3C XSLT is a language for transforming XML documents into other XML documents.[163]

XSLT is designed for use as part of XSL, which is a stylesheet language for XML. In addition to XSLT, XSL includes an XML vocabulary for specifying formatting. XSL specifies the styling of an XML document by using XSLT to describe how the document is transformed into another XML document that uses the formatting vocabulary.[164]

This specification can be used, for example, in the transformation from XSD documents defining GML documents to linked-open data formats (such as RDF, geoSPARQL, etc.).

7.5 Some Semantic Web tools for data meaning management

In this subsection, some W3C recommendations that are part of the W3C's Semantic Web technology stack and are of interest for data modelling and semantic information management are briefly described. A more exhaustive explanation of them is available by following the cited references.

7.5.1 Resource Description Framework (RDF)

RDF is an extension to XML and constitutes framework for representing information in the Web. The two key data structures are: RDF graphs, sets of subject-predicate-object triples, where the elements may be IRIs (Internationalized Resource Identifiers[165]*), blank nodes, or datatyped literals. They are used to express descriptions of resources. RDF datasets are used to organize collections of RDF graphs.*

The core structure of the abstract syntax is a set of triples, each consisting of a subject, a predicate and an object. A set of such triples is called an RDF graph. An RDF graph can be visualized as a node and directed-arc diagram, in which each triple is represented as a node-arc-node link (Figure 81). An RDF dataset is a collection of RDF graphs.[166]

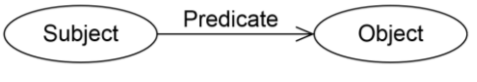


Figure 81. RDF graph: two nodes (Subject and Object) and a triple connecting them (Predicate)

[162] http://www.w3schools.com/xml/xml_xlink.asp
[163] https://www.w3.org/TR/xslt
[164] The italic text cites https://www.w3.org/TR/xslt
[165] https://www.w3.org/International/O-URL-and-ident.html
[166] The italic text cites https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/

7.5.2 RDF Schema

RDF Schema provides a data-modelling vocabulary for RDF data. RDF Schema is an extension of the basic RDF vocabulary.

RDF Schema is a semantic extension of RDF. It provides mechanisms for describing groups of related resources and the relationships between these resources. RDF Schema is written in RDF. These resources are used to determine characteristics of other resources, such as the domains and ranges of properties.

The RDF Schema class and property system is similar to the type systems of object-oriented programming languages: this is an advantage for the conceptual reasons explained in 2.2.2.5 and for being similar to existing and powerful programming paradigms.

RDF Schema differs from many such systems in that instead of defining a class in terms of the properties its instances may have, RDF Schema describes properties in terms of the classes of resource to which they apply. This is the role of the domain and range mechanisms described in this specification.

Using the RDF approach, it is easy for others to subsequently define additional properties with a domain or a range.

One benefit of the RDF property-centric approach is that it allows anyone to extend the description of existing resources, one of the architectural principles of the Web.

Resources may be divided into groups called classes. The members of a class are known as instances *of the class. Classes are themselves resources. They are often identified by IRIs and may be described using RDF properties. The rdf:type property may be used to state that a resource is an instance of a class.*

RDF distinguishes between a class and the set of its instances. Two classes may have the same set of instances but be different classes. A class may be a member of its own class extension and may be an instance of itself.

The group of resources that are RDF Schema classes is itself a class called `rdfs:Class`.

If a class C is a subclass *of a class C', then all instances of C will also be instances of C'. The rdfs:subClassOf property may be used to state that one class is a subclass of another. The term super-class is used as the inverse of subclass. If a class C' is a super-class of a class C, then all instances of C are also instances of C'.*

The RDF Concepts and Abstract Syntax specification defines the RDF concept of an RDF datatype. All datatypes are classes. The instances of a class that is a datatype are the members of the value space of the datatype.[167]

7.5.3 Web Ontology Language (OWL)

The W3C Web Ontology Language (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. OWL is a computational logic-based language such that knowledge expressed in OWL can be exploited by

[167] The italic text cites https://www.w3.org/TR/rdf-schema/

computer programs. OWL documents, known as ontologies, can be published in the World Wide Web and may refer to or be referred from other OWL ontologies.[168]

OWL provide a richer vocabulary or 'ontology' language, permitting to define inference rule languages and other formalisms, for example temporal logics and temporally changing information (Zamborlini, Guizzardi, 2010). These will contribute to our ability to capture meaningful generalizations about data in the Web.[169]

7.5.4 SPARQL Protocol and RDF Query Language (SPARQL)

RDF is a directed, labelled graph data format for representing information in the Web. This specification defines the syntax and semantics of the SPARQL query language for RDF. SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL contains capabilities for querying required and optional graph patterns along with their conjunctions and disjunctions. SPARQL also supports extensible value testing and constraining queries by source RDF graph. The results of SPARQL queries can be results sets or RDF graphs.

Most forms of SPARQL query contain a set of triple patterns called a basic graph pattern. *Triple patterns are like RDF triples except that each of the subject, predicate and object may be a variable. A basic graph pattern* matches *a subgraph of the RDF data when RDF terms from that subgraph may be substituted for the variables and the result is RDF graph equivalent to the subgraph.*

The most simple query consists of two parts: the `SELECT` *clause identifies the variables to appear in the query results, and the* `WHERE` *clause provides the basic graph pattern to match against the data graph, similarly to SQL.*[170]

7.6 OPEN STAND - The Modern Paradigm for Standards

Figure 82. Open stand logo

On 29 August 2012, five leading global organizations jointly signed an agreement to affirm and adhere to a set of Principles in support of The Modern Paradigm for Standards; an open and collectively empowering model that will help radically improve the way people around the world develop new technologies and innovate for humanity. Since then, hundreds of proponents from industry, civil

[168] The italic text cites https://www.w3.org/2001/sw/wiki/OWL
[169] The italic text cites https://www.w3.org/TR/rdf-schema/
[170] The italic text cites https://www.w3.org/TR/rdf-sparql-query/

society, government and academia, as well as individual technologists and innovators have expressed their support of the principles.

OpenStand is a movement dedicated to promoting a proven set of principles through which it encourages the development of market-driven standards that are global and open, enabling standards without borders and driving innovation for the benefit of humanity.

The OpenStand principles are based on the effective and efficient standardization processes that have made the Internet and Web the premiere platforms for innovation and borderless commerce and are extendable to other technologies. The principles stress voluntary adoption and empower the economies of global markets, fuelled by technological innovation, to drive global standards deployment. The result is The Modern Paradigm for Standards.

This approach resulted in the advancement of cutting-edge technology based on merit, empowering the rapid, economical implementation of high-value, high-demand products and services. The application of open principles resulted in more widespread acceptance of new standards within the global marketplace and drove more rapid development of the Internet and World Wide Web. The results were unprecedented, fuelling an economic and social transformation, and touching billions of lives.

The five stated principles are:

- *cooperation among standards organizations;*
- *adherence to the five fundamental principles of standards development (due process, broad consensus, transparency, balance and openness in standards development);*
- *commitment to technical merit, interoperability, competition, innovation and benefit to humanity;*
- *availability of standards to all;*
- *voluntary adoption.*[171]

7.7 An inherent issue: the open source community and the open source software

The philosophy (of which both Semantic Web and smart cities are products, promoters and boost) to share knowledge to achieve some personal and collective advantage, also involve the 'open source' world.

The 'open' concept refers to the transparency and readability of a technology. The obvious examples are open source software, which permits users to access their code and to change it or, in any case, to understand the algorthms that are employed. This allows the eventuality of participation of a wide, often voloutary, community for the improvement of the software or the involved technology. Private users write code for their own use and then share it to the collective, contributing to the improvement of the technology (Von Krogh et al., 2003).

Open source software originated around the 1950s-60s; then, in the 1980s (with the software being increasingly commercialized) Richard Stallmann, a researcher at MIT, founded the Free Software

[171] The italic text cites https://open-stand.org/about-us/principles/

Foundation, providing a conceptual foundation for open source software in the 'GNU manifesto' (GNU's Not Unix) (Stallman, 1985).

The history of open source software (OSS) has evolved up to the present (Hars, Ou, 2001). Today, the fact that open source often implements international official standards to maximize interoperability is a further challenge for the reasons previously described. Also, for this fact, OSS are currently spreading. Also, some laws are being enacted that encourage their use: for example, the Italian Circular 63/2013 (Agenzia per l'Italia Digitale, 2013) establishes the necessity for public administrations to prefer open source software.

OSS have some advantages compared with proprietary software. First of all, they permit the saving of expensive licensing and maintenance costs; moreover, they promote data interoperability by using open formats and implementing international standards. Finally, the knowledge of the source code is a key element to effectively monitor and control the operations of the software. This enables the user to adapt the processing to his specific needs and, where necessary, to correct, change or enhance it, contributing to its improvement.

However, OSS sometimes presents some difficulties to beginners or new users, since this form of software is often less user-friendly than proprietary software. Moreover, beta versions are often available and it is possible that algorithms or functionalities are not perfectly optimized and operating in those versions. Often, particular precautions or tricks are necessary in order to achieve results without errors. However, this typically does not affect the quality and efficiency of the final results.

For these reasons, when possible, OSS was used for the processing of this research.

7.8 Domain ontologies in the Semantic Web

Several ontologies exist that employ technologies, schemas, and tools as explained in Section 7.5. They are ontologies that deal with concepts or entities, which could be considered for some aspects of the study of architectural heritage and its context, or which should be related to the entities and resources defined in the application studied in this research.

The more relevant ontology for cultural heritage is the CIDOC-CRM, briefly presented in the next Section 7.8.1. Then, other projects or inventories that use it, are described in Section 7.8.2. Finally, other ontologies based on OWL definitions and connected to the management of architecture are reported in Sections 7.8.3 and 7.8.4.

7.8.1 The CIDOC CRM and its extensions: the main reference for cultural heritage semantics

The CRM Special Interest Group (LeBoeuf et al., 2013) of the International Committee for Documentation (CIDOC) of the International Council of Museums (ICOM) has defined the most important Conceptual Reference Model (CRM) for the information about cultural heritage (Doerr, 2009; 2002) since 1996. It became the International Standard ISO21127:2006 in September 2006 (ISO, 2006) and is now available as an OWL ontology[172]

[172] http://www.cidoc-crm.org/official_release_cidoc.html

The aim of this formal ontology was to enable the exchange and integration of information between heterogeneous sources of CH data. Originally it was particularly intended for the representation of the knowledge of museum objects; therefore, if used for different aims (architectures, archaeological heritage, etc.), it had to be adapted or expanded (De Roo et al., 2013). CIDOC-CRM can be regarded as a core ontology, as it models space-time concepts, events, and material and immaterial objects originating from several fields of study self-intersecting in the cultural heritage domain. It can be extended for being improved or applied to various, more specific application domains. Also, a management of time is included in the model, which can be managed as semantic data in RDF[173]

It surely can serve to build new repositories and to be a reference for mapping from existing databases for the aim of interoperability (Kondylakis et al., 2006). Furthermore, some general extensions of the model for permitting an improved exploitation of its application potentiality were developed. In particular, **CRMsci** is a more detailed model and extension of the CIDOC CRM to document scientific activity: it includes the representation of metadata about scientific observation, measurements, and processed data in descriptive and empirical sciences (Doerr et al., 2014). Another one is **CRMinf**, an extension to support argumentation and inference (Stead et al., 2015[174]).

In the **CRMgeo** extension of the CIDOC CRM (Doerr et al., 2013), OGC geoSPARQL (Section 7.9.1) was integrated to enable the model to represent more detail in the spatial component of its content and more complexity and articulation in time expressions.

The main lack in this ontology concerned the management of complex 3D spatial resources: in applications, often only the geographic position can be inserted and linked to some sort of webGIS to be visualized on the map (Hernàndez et al., 2008; the Getty ARCHES project – Section 7.8.2.2). Also, in CRMgeo a middle scale 2D space is usually referenced.

Another ontology developed in OWL and conceived as an enhancement of CIDOC-CRM is **MONDIS** (Monument Damage Information System) (Blaško et al., 2012, Cacciotti et al., 2013). This schema is developed for representing the information about preservation, decay, deterioration mechanisms, restoration and intervention.

An extension of particular interest in the framework of this thesis is the recently developed **CRMBA**, expressly realized for the documentation of standing buildings (Ronzino et al., 2015). This last extension has almost identical aims of this thesis and the spatial issues are managed with the other extension of the CIDOC CRM that is of the very interest of this thesis: CRMgeo. The only gap in this research with respect to the considered issues in this thesis could be found in the management of 3D models with built structures and in connection with other parts of the city and the landscape, which is a topic treated by CityGML. An integration between the two will be pursued in the future.

7.8.2 Some projects employing the CIDOC-CRM

7.8.2.1 EPOCH (Excellence in Processing Open Cultural Heritage)

EPOCH is the EU FP6 Network of Excellence on the Applications of ICT (Information and Communication Technologies) to tangible cultural heritage, realized in cooperation with 95 European partners. The project was active from 2004 to 2008. Among the project's subfields were field recording and data capture, data organisation provenance and standards, and reconstruction and visualisation.

[173] http://www.cidoc-crm.org/docs/How_to%20implement%20CRM_Time_in%20RDF.pdf
[174] http://www.ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf0.5.pdf

The aim was to build an inter-disciplinary structure of data about cultural heritage exploiting Information and Communication Technologies and open principles in order to promote the knowledge, communication and promotion of cultural heritage with a holistic view. In the project EPOCH the employment of open instruments to generate Open Digital Cultural Heritage is one of the objectives.

Perhaps some technologies were not yet sufficiently mature to realize some goals pursued in the project, but important concepts are affirmed. Among these: the need for interoperability through common standards, open formats, and open procedures and tools; the need for conferring semantic meaning to 3D digital artefacts for a correct interpretation; and the need for organising the thematic data regarding artefacts in databases related to 3D products.[175]

For realizing this semantic enrichment of the resources, enabling their integration, query, and reuse, the CIDOC CRM has been used as reference[176]. The necessity was to map the data from various known or unknown encodings to the chosen ontology. For this reason, the **AMA** project was developed (Archive Mapper for Archaeology). It is an open tool for the semi-automated mapping of archaeological data to CIDOC CRM[177] (Eide et al., 2008).

A further tool developed to semantically manage archaeological data was **MAD** (Managing Archaeological Data)[178]. It is an application designed to store, manage, and browse structured and unstructured archaeological datasets encoded in a semantic format. Also, this tool is based on Open Source technologies, XML, and W3C standards, and it uses the CIDOC-CRM and future developments to include geometrical data management in order to manage stratigraphy and geographic information as well.

As matter of fact standardization is still far to achieve in Cultural Heritage, and EPOCH is committed to remove any obstacle that may delay this objective. The Network fosters the application of the CIDOC Conceptual Reference Model-CRM (ISO 21127) at all its partners, facilitating this through training dissemination and access to documentation. Mapping partners' internal data structures to CIDOC-CRM is also a way to promote the adoption of this international standard. Concerning technological standards, EPOCH is actively investigating those more suitable for use in 3D applications to cultural heritage.

EPOCH strongly supports the establishment of charters concerning the communication of Cultural Heritage, such as the Ename Charter […]. EPOCH is also pushing forward the definition of an international charter on the credibility of virtual reconstructions, named the London Charter.[179]

7.8.2.2 The ARCHES project

The ARCHES project is a collaboration between the Getty Conservation Institute (GCI) (4.1) *and World Monuments Fund (WMF) to develop for the international heritage field an open source, web- and geospatially based information system that is purpose-built to inventory and manage immovable cultural heritage. ARCHES incorporates widely adopted standards (for heritage inventories, heritage data and information technology) so that it will offer a solid foundation that heritage institutions may customize to meet their particular needs. ARCHES is built using open source software tools to make*

175 http://public-repository.epoch-net.org/reports/CI_Tool_interpretation_management_v7.pdf

176 http://public-repository.epoch-net.org/presentations/rome/Procedures%20for%20Data%20Integration%20through%20CIDOC%20CRM.pdf

177 http://epoch-net.org/site/research/newtons/ama/

178 http://epoch-net.org/site/mad/

179 The italic text cites http://epoch-net.org/site/research/standards/

its adoption cost effective, and to allow heritage institutions to pool resources to enhance ARCHES in mutually beneficial ways.[180]

The ARCHES project team determined that, based on these efforts and experience developing the MEGA system, ARCHES should be designed according to the following principles:

- ***standards based:** ARCHES is being built using internationally adopted standards in the cultural heritage and information technology fields;*
- ***broadly accessible:** ARCHES will be web-based to provide for the widest possible access. It will be user friendly and require minimal training for most users. The system will be freely available for download from the Internet so that institutions may customize it and install it at any location in the world;*
- ***economical to adapt and implement:** Through the use of open source tools, it will be economical to customize, update, and maintain and will be vendor neutral with no licensing or upgrade fees;*
- ***customizable:** The software code will be open sourced, allowing it to be readily customized. The system will be structured in modules so that it may be easily extended. It will be capable of presenting its user interface in any language, or in multiple languages, including those that are bidirectional and use non-Latin scripts. It will be configurable to any geographic location or region;*
- ***secure:** Once the Arches system is installed, institutions implementing it may control the degree of privacy of their data that the system contains.*

Arches has been designed to serve a number of purposes fundamental to the understanding, appreciation, and management of heritage places. These include: identification and inventory, research and analysis, monitoring and risk mapping, determining needs and priorities for investigation and research, conservation and management, planning for investigation, conservation, and management activities, raising awareness and promoting understanding among the public, as well as governmental authorities and decision makers.

ARCHES also incorporates international heritage documentation standards. The International Core Data Standard for Archaeological and Architectural Heritage, the combined standards that is being finalized by the CIDOC (6.2.4), was used to identify the data fields of version 1.0. Organizations that deploy ARCHES can customize those data fields to meet their specific requirements. ARCHES bases the relationships between data fields in the system on the CIDOC CRM - ISO 21127:2006. Use of the CRM keeps the data independent of conventions that are particular to the design of Arches. It also contributes to powerfully effective searches within, as well as across, data sets. It will facilitate data migration to newer systems and aid in the preservation of data over time.[181]

The principles followed in the project state interoperability sharing and accessibility, which are common to this research. A webGIS is included in the inventory: it permits to draw a geometry (point / line / polygon) collocating the item on a base map. A complex management of geometry is however missing.

7.8.2.3 The Getty Vocabularies

The Getty vocabularies contain structured multilingual terminology for art, architecture, decorative arts, archival materials, visual surrogates, conservation, and bibliographic materials. Compliant with

[180] The italic text cites http://www.getty.edu/conservation/our_projects/field_projects/arches/arches_overview.html
[181] The italic text cites http://www.getty.edu/conservation/our_projects/field_projects/arches/arches_overview.html

international standards, they provide authoritative information for cataloguers, researchers, and data providers. The vocabularies grow through contributions. In the new linked, open environments, they provide a powerful conduit for research and discovery for digital art history.

Getty vocabularies are thesauri compliant with national (NISO) and international (ISO) standards for thesaurus construction. They are compliant with / map to other standards e.g.:

- *CDWA (Categories for the Description of Works of Art)*
- *CCO (Cataloging Cultural Objects)*
- *VRA Core (Visual Resources Association core categories)*
- *LIDO (Lightweight Information Describing Objects)*
- *CIDOC CRM (CIDOC Conceptual Reference Model) (ISO 21127:2006)*[182]

The Getty vocabularies (Harpring, 2010) are built to allow their use in linked data[183]. A project to publish them to the LOD (Linked Open Data) cloud is underway[184]. For doing the passage, URIs (Uniform Resource Identifiers) for identifying the resources were defined. They are based on the unique, persistent subject_ID of the vocabularies.

Example of URIs from each vocabulary:

- Human-readable full record:

http://vocab.getty.edu/page/aat [acronym of the vocabulary] /300198841 [resource subject_ID]

- Human-readable hierarchy view:

http://vocab.getty.edu/hier/aat [acronym of the vocabulary] /300198841 [resource subject_ID]

- Semantic RDF concept:

http://vocab.getty.edu/aat [acronym of the vocabulary] /300198841 [resource subject_ID]

7.8.2.3.1 The Art & Architecture Thesaurus® (AAT)

This Getty vocabulary includes generic terms for describing art and architecture (decorative arts, other material culture, visual surrogates, archival materials, archaeology, and conservation).[185] A research example is reported in Figure 83 - Figure 85.

Current totals: 36000 records; 245000 terms

AAT is multilingual, and large translation projects are underway. Conceptually organized terms about both abstract concepts and generic terms for concrete, physical artefacts are included.

[182] The italic text cites http://vocab.getty.edu
[183] http://www.getty.edu/research/tools/ vocabularies/Linked_Data_Getty_Vocabularies.pdf
[184] http://www.getty.edu/research/tools/vocabularies/lod/index.html
[185] http://www.getty.edu/research/tools/vocabularies/aat/index.html

Research

Research Home ▸ Tools ▸ Art & Architecture Thesaurus ▸ Full Record Display

Art & Architecture Thesaurus® Online
Full Record Display

New Search   ◂ Previous Page   Help

Click the icon to view the hierarchy.

Semantic View (JSON, JSONLD, RDF, N3/Turtle, N-Triples)

ID: 300001662   **Record Type:** concept

capitals (column components) (<capitals and capital components>, column components, ... Components (hierarchy name))

Note: The uppermost members of columns, piers, or pilasters.

Terms:

capitals (column components) (**preferred**,C,U,LC,English-P,D,U,PN)
capital (column component) (C,U,English,AD,U,SN)
chapiters (C,U,English,UF,U,N)
柱頭 (立柱構造) (C,U,Chinese (traditional)-P,D,U,U)
柱頂 (C,U,Chinese (traditional),UF,U,U)
zhù tóu (C,U,Chinese (transliterated Hanyu Pinyin)-P,UF,U,U)
zhu tou (C,U,Chinese (transliterated Pinyin without tones)-P,UF,U,U)
chu t'ou (C,U,Chinese (transliterated Wade-Giles)-P,UF,U,U)
kapitelen (C,U,Dutch-P,D,U,U)
kapiteel (C,U,Dutch,AD,U,U)
capitelli (C,U,Italian-P,D,U,PN)
capitello (C,U,Italian,AD,U,SN)
capiteles (C,U,Spanish-P,D,U,PN)
capitel (C,U,Spanish,AD,U,SN)
chapitel (C,U,Spanish,UF,U,SN)

Facet/Hierarchy Code: V.PJ

Hierarchical Position:

Objects Facet
.... Components (hierarchy name) (G)
.... components (objects parts) (G)
.... <components by specific context> (G)
.... architectural elements (G)
.... <structural elements and components for structural elements> (G)
.... structural elements (G)
.... <supporting and resisting elements> (G)
.... <columns and column components> (G)
.... column components (G)
.... <capitals and capital components> (G)
.... capitals (column components) (G)

Figure 83. *Example of research results in AAT vocabulary - I.*

Sources and Contributors:

capital (column component)............ [VP]
.... Getty Vocabulary Program rules
capitals (column components)............ [GCI Preferred, VP Preferred]
.... AATA database (2002-) 121487 checked 26 January 2012
.... Avery Index (1963-) Capitals (Architecture)
.... Boethius, Etruscan and Early Roman Architecture (1978) GLOS
.... CDMARC Subjects: LCSH (1988-) Capitals (Architecture)
.... Dinsmoor, Architecture of Ancient Greece (1975) GLOS
.... Harris, Dictionary of Architecture and Construction (1975) illustration
.... Jacoby and Talgam, Jerusalem Index of Jewish Art (1988) illustration, 17
.... Nomenclature 3.0 for Museum Cataloging (2010)
.... RIBA, Architectural Keywords (1982) Capitals (columns)
.... RILA, Subject Headings (1975-1990)
capitel............ [CDBP-DIBAM Preferred]
.... Putnam y Carlson, Diccionario de Arquitectura, Construcción y Obras Públicas (1998) 456
.... Pérez Montás, La platería de la Catedral de Santo Domingo, primada de América (1993) 65
capiteles............ [CDBP-DIBAM]
.... Comité, Plural del término en singular
capitelli............ [VP]
.... Getty Vocabulary Program rules
capitello............ [VP]
.... Dizionario enciclopedico di architettura e di urbanistica (1968-1969) 1: 479
chapitel............ [CDBP-DIBAM]
.... Alvarez et al., Diccionario de la decoración (1968) 199
chapiters............ [VP]
.... Harris, Historic Architecture Sourcebook (1977)
.... Webster's Third New International Dictionary (1961)
chu t'ou............ [TELDAP-now AS]
.... TELDAP database (2009-)
kapiteel............ [RKD, AAT-Ned]
.... AAT-Ned (1994-)
kapitelen............ [RKD, AAT-Ned Preferred]
.... AAT-Ned (1994-)
zhu tou............ [TELDAP-now AS]
.... TELDAP database (2009-)
zhù tóu............ [TELDAP-now AS]
.... TELDAP database (2009-)
柱頭 (立柱構造)............ [TELDAP-now AS Preferred]
.... 世界建築經典圖鑑 p. 336
.... 世界藝術史 p. 890
.... 劍橋藝術史－希臘與羅馬 p. 131
.... 大英視覺藝術百科全書 Vol. 9,pp. 173
.... 朗文當代大辭典 p. 243
.... 朗文英漢雙解科技大辭典 p. 243
.... 牛津當代大辭典 p. 243
柱頂............ [TELDAP-now AS]
.... 劍橋藝術史－中世紀 p. 122

Figure 84. Example of research results in AAT vocabulary - II.

Subject: ..... [CDBP-DIBAM, GCI, RKD, AAT-Ned, TELDAP-now AS, VP]
.... AATA database (2002-) 121487 checked 26 January 2012
.... Avery Index (1963-) Capitals (Architecture)--Doric; Capitals (Architecture)--Ionic; Capitals (Architecture)--Islamic; (source AAT); Capitals (Architecture)--Composite; Capitals (Architecture)--Corinthian
.... Haslinghuis, Woordenboek bouwkundige termen (1986)
.... RIBA, Architectural Keywords (1982) Columns: capitals
.... RILA, Subject Headings (1975-1990) precoor.; Roman capitals; Coptic capitals; Gothic capitals; Islamic capitals; Medieval capitals; Crusader capitals; Byzantine capitals; Romanesque capitals; Carolingian capitals; Renaissance capitals; Merovingian capitals; Anglo-Saxon capitals; Late Antique capitals; Early Christian capitals

Note:
English .... [VP]
.... Dizionario enciclopedico di architettura e di urbanistica (1968-1969)
Chinese (traditional).......... [TELDAP-now AS]
.... TELDAP database (2009-)
Dutch .... [RKD, AAT-Ned]
.... AAT-Ned (1994-)
.... Haslinghuis, Bouwkundige termen (1986)

---

w Search

ck to top

inter-friendly version

The J. Paul Getty Trust

Figure 85. Example of research results in AAT vocabulary - III.

7.8.2.3.2 *The Getty Thesaurus of Geographic Names® (TGN)*

This deals with names, other information for current and historical administrative places (cities, nations, empires) and physical features (e.g., Firenze, Roman Empire, Ganges River).[186]

Current totals: 1241000; 1800000 names

The TGN vocabulary includes formerly inhabited places, archaeological sites, and historical places with unknown exact locations (i.e., 'lost settlements'). It focuses on places important to the study of art and related disciplines. This resource grows through contributions from large national geographic databases, the expert user community, archaeology projects, scholars, etc.

It is a useful integration to the GeoNames database when one needs to manage data about historical artefacts. It constitutes a sort of historic gazetteer, with some synthetic information added to the list of names in various languages and changed over the time. An example of research result is shown in Figure 86 - Figure 87.

Figure 86. Example of research result in the Getty TGN - I

[186] http://www.getty.edu/research/tools/vocabularies/tgn/index.html

Additional Parents:

World (facet)
.... Europe (continent) (P)
.... Savoy (historical region) (P,H)
.... Pinerolo (inhabited place) (P)

Place Types:

inhabited place (**preferred**, C) .... first documented in 996
town (C)
transportation center (C) .... is a rail junction
manufacturing center (C) .... for textiles, metal, chemicals, printing & food products
episcopal see (C) .... established in 1748
prison center (H) .... fortress was state prison for 17th-cen. enemies of Louis XIV, including "Man in Iron Mask" (as in novel by A Dumas père)
noble seat (H) .... seat of the counts of Acaia, a subsidiary line of the House of Savoy, 1295-1418

Sources and Contributors:

Pignerol.......... [GRLPSC]
.... Webster's Geographical Dictionary (1984)
Pinerolo.......... [BHA, GRLPSC, VP Preferred]
.... Annuario Generale (1980)
.... Columbia Lippincott Gazetteer (1961)
.... Encyclopaedia Britannica (1988) IX, 454
.... NGA/NIMA database (2003-)
.... Times Atlas of the World. Reprint ed. (1994) 151
.... Webster's Geographical Dictionary (1984)
Subject: ..... [BHA, GRLPSC, VP]
.... Annuario Generale (1980)
.... Columbia Lippincott Gazetteer (1961)
.... Encyclopaedia Britannica (1988) IX, 454
.... NGA/NIMA database (2003-) -124863
.... Times Atlas of the World. Reprint ed. (1994) 151
.... Webster's Geographical Dictionary (1984)

Note:
English.......... [VP]

New Search

Back to top

Printer-friendly version

The J. Paul Getty Trust

© 2004 J. Paul Getty Trust
Terms of Use / Privacy Policy / Contact Us

Figure 87. Example of research result in the Getty TGN – II

7.8.2.3.3 The Union List of Artist Names® (ULAN)

This vocabulary collects names and other information for people (artists, patrons, sitters) and corporate bodies (repositories, studios) related to the design, creation, patronage, collection, conservation, etc. of art, architecture, and other cultural materials. It includes current and historical associative relationships (e.g., student-teacher, firm-member).[187] An example of a research result is shown in Figure 88 - Figure 90.

Current totals: 223000 records; 582000 names

Contributions to ULAN are from expert communities (museums, art libraries, cataloguing projects, etc.). ULAN is contributed to the Virtual International Authority File (VIAF) (a joint project with the US Library of Congress and numerous libraries worldwide to combine name authority).

Research

Research Home ▸ Tools ▸ Union List of Artist Names ▸ Full Record Display

Union List of Artist Names® Online
Full Record Display

New Search ◂Previous Page Help

Click the icon to view the hierarchy.

Semantic View (JSON, JSONLD, RDF, N3/Turtle, N-Triples) ▪ Representative Images: 1

ID: 500022389 **Record Type:** Person

Juvarra, Filippo (Italian architect and scenographer, 1678-1736, active in Italy and Spain)

Note: Important and influential Baroque architect who introduced elements of the Rococo architectural style. After working in his father's silversmith shop and being ordained a priest in 1703, Juvarra began his architectural training at around age twenty-five. After moving to Rome, he studied under its leading Baroque architect and gained renown for his bold stage designs. He was named First Architect to the King of Sicily, and charged with transforming Turin, Savoy's capital. He produced sixteen palaces, eight churches, urban planning projects, interior decoration, and designs for furniture and church ornaments. Highly skilled painters, sculptors, and craftsmen from throughout Italy made his designs reality. In 1735 Philip V summoned Juvarra for work in Spain, where his sudden death cut short his career.

Names:
- **Juvarra, Filippo** (**preferred**,V,index)
- **Filippo Juvarra** (V,display)
- **Hovara, Filippo** (V)
- **Ivara, Filippo** (V)
- **Ivvara, Filippo** (V)
- **Juara, Filippo** (V)
- **Juarra, Filippo** (V)
- **Juvara, Filippo** (V,LC)
- **Filippo Juvara** (V)

Nationalities:
- Italian (**preferred**)
- Spanish
- Sicilian

Roles:
- artist (**preferred**)
- architect
- printmaker
- etcher
- scenographer
- silversmith
- draftsman
- designer

Figure 88. Example of research result in the Getty ULAN - I

[187] http://www.getty.edu/research/tools/vocabularies/ulan/index.html

Gender: male

Birth and Death Places:
Born: Messina (Messina province, Sicily, Italy) (inhabited place)
Died: Madrid (Madrid province, Comunidad de Madrid, Spain) (inhabited place)

Events:
active: Madrid (Madrid province, Comunidad de Madrid, Spain) (inhabited place)
active: Rome (Roma province, Lazio, Italy) (inhabited place)
active: Turin (Torino province, Piedmont, Italy) (inhabited place)

Related People or Corporate Bodies:
assisted by .... Sacchetti, Giovanni Battista
.... (Italian architect, 1690-1764, active in Spain) [500012096]
influenced .... Vittone, Bernardo Antonio
.... (Italian architect and writer, 1702-1770) [500008023]
patron was .... Elizabeth Farnese, Queen consort of Philip V
.... (Spanish queen consort, 1692-1766) [500039510]
patron was .... Philip V, King of Spain
.... (Spanish king, 1683-1746) [500122373]
student of .... Fontana, Carlo
.... (Italian architect and engineer, 1634-1714) [500026624]
uncle/aunt of .... Passalacqua, Pietro
.... (Italian architect, died 1748) [500001821]

List/Hierarchical Position:
.... Persons, Artists
.... Juvarra, Filippo (I)

Biographies:
(Italian architect and scenographer, 1678-1736, active in Italy and Spain) ..... [VP Preferred]
(artist, 1678-1736) ..... [GRL]
(Italian architect, 1678-1736) ..... [Grove Art]
(Italian artist, 1674/1676-1736) ..... [PROV]
(architect, 1678-1736) ..... [CL-Courtauld]
(Italian architect, draftsman and scenographer, 1678-1736) ..... [VP]
(Italian architect) ..... [AVERY]
(Italian architect, 1678-1736) ..... [AVERY]
(Italian artist, 1674/1676-1736) ..... [WL-Courtauld]
(1678-31.I.1736; Late Baroque Classical Architect, Scenography Designer, Madrid, Roma, Torino) ..... [FDA]
(Italian artist, 1674/6-1736) ..... [WCP]
(Italian architect and printmaker, 1676-1736) ..... [GRLPSC]
(Italian architect, etcher and theatre designer; born Messina (Italy), 1678; died Madrid (Spain), 1736) ..... [CCA]
(Italian architect, printmaker, 1678-1736) ..... [BHA]
(Italian architect, 1678-1736) ..... [AVERY]

Sources and Contributors:
Filippo Juvara .... [CL-Courtauld, PROV, VP]
.... Getty Vocabulary Program rules
Filippo Juvarra .... [CCA, VP]
.... Canadian Centre for Architecture database
Hovara, Filippo .... [AVERY, CCA, GRL]
.... Avery Authority files (1963-)

Figure 89. Example of research result in the Getty ULAN - II

.... Avery Index to Architectural Periodicals (1963-)
.... Library of Congress Authorities database (n.d.) n 79127059
Ivara, Filippo .... [CCA, FDA, GRL, WL-Courtauld]
.... Bénézit, Dictionnaire des Peintres (1976)
.... Database for the Witt Checklist of Painters c. 1200-1976 (1978-)
.... Library of Congress Authorities database (n.d.) n 79127059
.... Witt Library, Authority files
Ivvara, Filippo .... [AVERY]
.... Avery Authority files (1963-)
.... Avery Index to Architectural Periodicals (1963-)
Juara, Filippo .... [CCA, FDA]
.... Bénézit, Dictionnaire des Peintres (1976)
Juarra, Filippo .... [CCA, GRL, WL-Courtauld]
.... Canadian Centre for Architecture database
.... Database for the Witt Checklist of Painters c. 1200-1976 (1978-)
.... Library of Congress Authorities database (n.d.) n 79127059
.... Witt Library, Authority files
Juvara, Filippo .... [AVERY Preferred, BHA Preferred, CCA Preferred, CL-Courtauld Preferred, FDA, GRL Preferred, GRLPSC Preferred, PROV Preferred, WL-Courtauld Preferred]
.... Avery Authority files (1963-)
.... Database for the Witt Checklist of Painters c. 1200-1976 (1978-)
.... Library of Congress Authorities database (n.d.) n 79127059
.... Library of Congress Authorities online (2002-)
.... Provenance Index Databases, Authority file (1985-)
.... RILA/BHA (1975-2000) 1986
.... Witt Library, Authority files
Juvarra, Filippo .... [AVERY, CCA, FDA Preferred, GRL, Grove Art Preferred, JPGM Preferred, VP Preferred, WL-Courtauld]
.... Avery Authority files (1963-)
.... Database for the Witt Checklist of Painters c. 1200-1976 (1978-)
.... Grove Art artist database (1989-)
.... J. Paul Getty Museum, collections online (2000-)
.... Library of Congress Authorities database (n.d.) n 79127059
.... Macmillan Encyclopedia of Architects (1982)
.... National Gallery of Art database, Washington (1987-)
.... Witt Library, Authority files

Subject: .... [AVERY, BHA, CCA, CL-Courtauld, FDA, Gallery Systems, GRL, GRLPSC, Grove Art, JPGM, PROV, VP, WCP, WL-Courtauld]
.... Allgemeines Künstlerlexikon. Internationale Künstlerdatenbank (1993-2006)
.... Avery Index to Architectural Periodicals (1963-)
.... Bessone-Aurelj, Dizionario scultori ed architetti (1947)
.... Bolaffi, Dizionario dei pittori italiani (1972-1976)
.... Bénézit, Dictionnaire des Peintres (1976)
.... Canadian Centre for Architecture database
.... Database for the Witt Checklist of Painters c. 1200-1976 (1978-)
.... Enciclopedia Italiana di Scienze, Lettere ed Arti (1949-1951)
.... Gallery Systems (2000-) LOC ID: n79127059
.... Getty Vocabulary Program rules
.... Grove Art artist database (1989-)
.... J. Paul Getty Museum, collections online (2000-)
.... LCNAF Library of Congress Name Authority File [n.d.]
.... Library of Congress Authorities database (n.d.) LOC ID: n79127059
.... Library of Congress Authorities online (2002-)
.... Macmillan Encyclopedia of Architects (1982)
.... Millon, Filippo Juvarra (1984)
.... National Gallery of Art database, Washington (1987-)
.... RILA/BHA (1975-2000) 1986
.... Thieme-Becker, Allgemeines Lexikon der Künstler (1980-1986) v.19, p.358 ff.
.... Witt Library, Authority files

Note:
English.......... [AVERY, VP]
.... J. Paul Getty Museum, collections online (2000-) accessed 5 November 2008
.... LCNAF Library of Congress Name Authority File [n.d.]

New Search

Back to top

Printer-friendly version

The J. Paul Getty Trust

© 2004 J. Paul Getty Trust
Terms of Use / Privacy Policy / Contact Us

Figure 90. Example of research result in the Getty ULAN - IV

7.8.2.3.4 The Cultural Objects Name Authority® (CONA)

This vocabulary includes titles, other information for moveable works (paintings, sculpture, prints, drawings, photographs, ceramics, etc.) and architecture (buildings, bridges, etc.) (e.g., Hagia Sophia, Mona Lisa, Fantastic Landscape with a Pavilion). It includes works executed or designed (e.g., a built work for which only design drawings exist), extant or destroyed/historical.[188] Its terms can be linked to the other vocabularies and to other CONA records.

Current totals: 5500 records; 9300 titles

CONA is a thesaurus; it is also mapped to CDWA, CCO, CIDOC CRM, and LIDO. CONA grows through contributions from repositories of art and the expert cataloguing and scholarly community.

Under discussion are the possible future applications of CONA:

- as a resource to link the Getty vocabularies to records for art and architecture (in LOD or otherwise);
- to link subject terminology to art and architecture records;
- to facilitate linking between works, including works held in different repositories but having historical relationships (e.g., studies, disassembled manuscripts, etc.).

Moreover, it could also serve as a reference for linking 3D semantic spatial models.

Figure 91, Figure 92, Figure 93 show some examples of functions of CONA vocabulary. An example of a research result is shown in Figure 94 - Figure 95.

Figure 91. Example of resources with similar subjects that could be related in CONA (Harpring, 2014[189])

[188] http://www.getty.edu/research/tools/vocabularies/cona/index.html
[189] https://www.getty.edu/research/tools/vocabularies/cidoc_getty_vocabs_lod.pdf

Figure 92. Example of resources representing parts held in different repositories belonging to the same object that can be related in CONA (Harpring, 2014). It is the proposed reconstruction of Matteo di Giovanni's Asciano Altarpiece by Rachel Billinge. Eextant panels are in the National Gallery in London, Siena, Asciano, Villa I Tatti, Esztergom, Altenburg, Rhode Island and private collections (Harpring, 2014). The same could happen, for example, with furniture and buildings, including the 3D model in the different locations, or with parts of previous monuments that have been re-used in the following buildings.

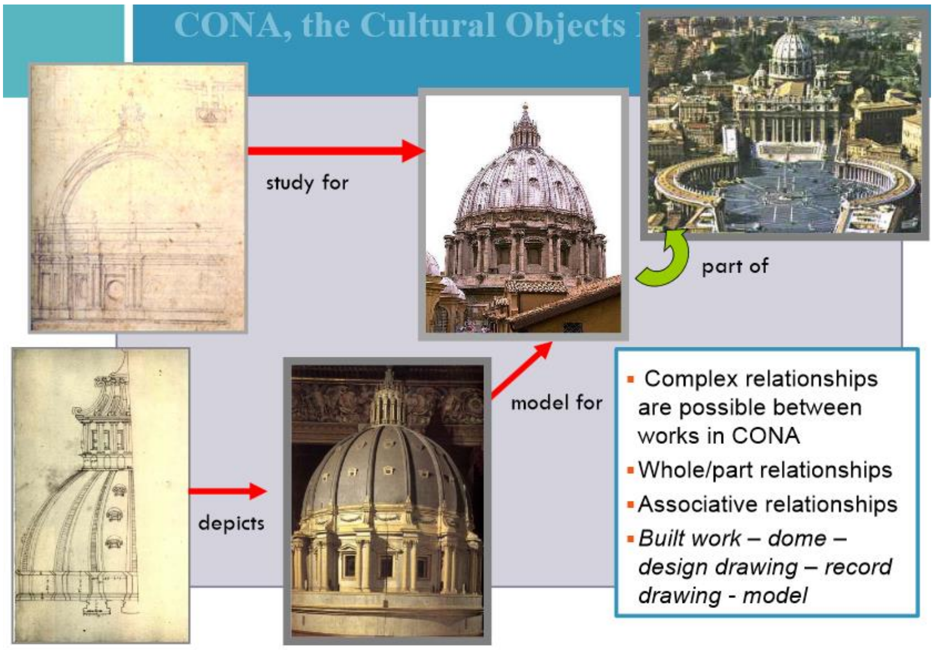
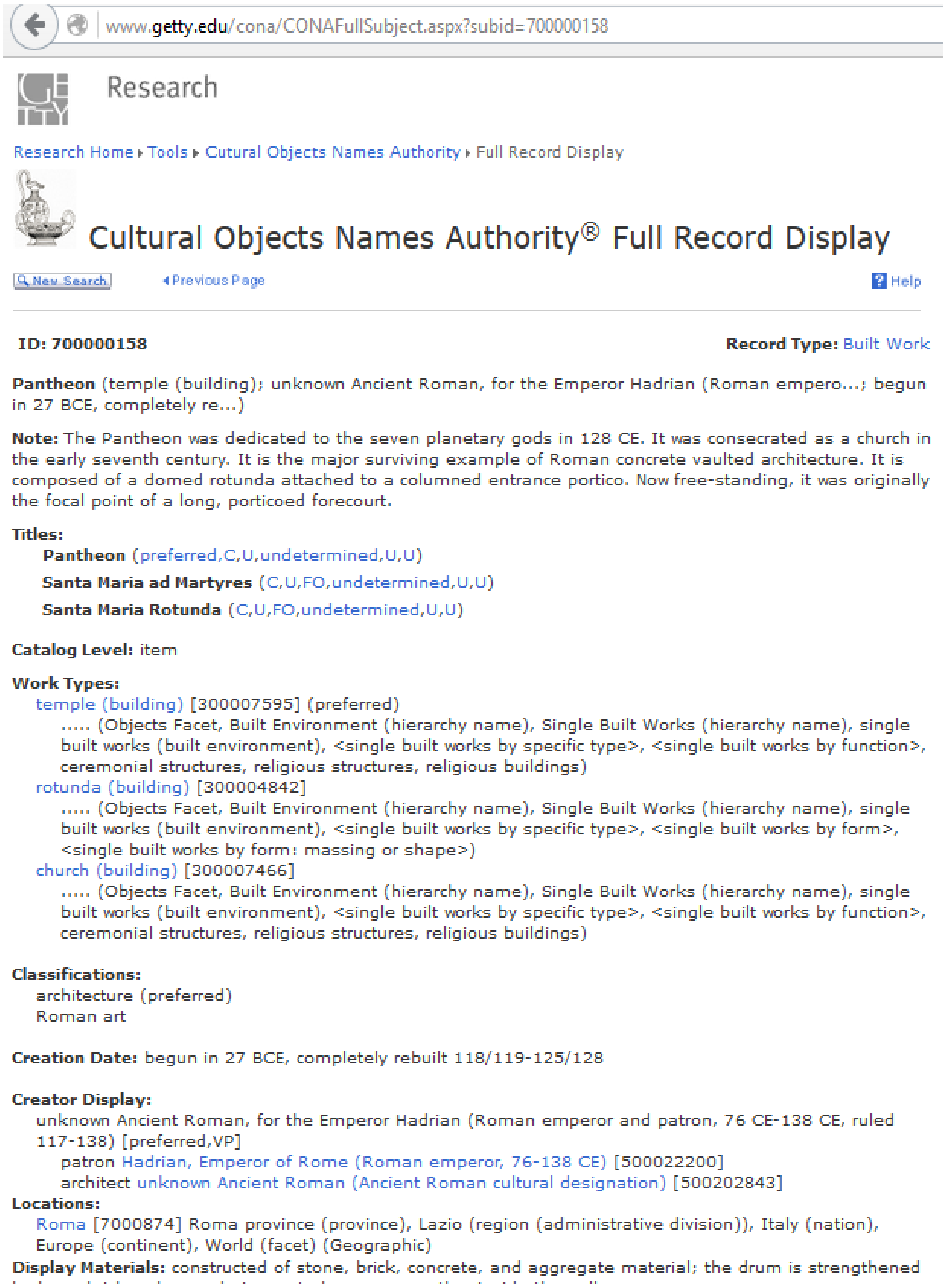


Figure 93. Another example of a possible relationship in CONA (Harpring, 2014). A number of projects and representation of a monument, held in different and possibly dispersed archives and publications or private repositories, can be stored and be available at the same time through the internet using LODs. The 3D model representation and all the geographic resources included in the built multi-representation database can be included for an integrated archiving of the information.



www.getty.edu/cona/CONAFullSubject.aspx?subid=700000158

Research

Research Home ▸ Tools ▸ Cutural Objects Names Authority ▸ Full Record Display

Cultural Objects Names Authority® Full Record Display

New Search   ◂Previous Page   Help

ID: 700000158   **Record Type:** Built Work

Pantheon (temple (building); unknown Ancient Roman, for the Emperor Hadrian (Roman empero...; begun in 27 BCE, completely re...)

Note: The Pantheon was dedicated to the seven planetary gods in 128 CE. It was consecrated as a church in the early seventh century. It is the major surviving example of Roman concrete vaulted architecture. It is composed of a domed rotunda attached to a columned entrance portico. Now free-standing, it was originally the focal point of a long, porticoed forecourt.

Titles:
Pantheon (preferred,C,U,undetermined,U,U)
Santa Maria ad Martyres (C,U,FO,undetermined,U,U)
Santa Maria Rotunda (C,U,FO,undetermined,U,U)

Catalog Level: item

Work Types:
temple (building) [300007595] (preferred)
..... (Objects Facet, Built Environment (hierarchy name), Single Built Works (hierarchy name), single built works (built environment), <single built works by specific type>, <single built works by function>, ceremonial structures, religious structures, religious buildings)
rotunda (building) [300004842]
..... (Objects Facet, Built Environment (hierarchy name), Single Built Works (hierarchy name), single built works (built environment), <single built works by specific type>, <single built works by form>, <single built works by form: massing or shape>)
church (building) [300007466]
..... (Objects Facet, Built Environment (hierarchy name), Single Built Works (hierarchy name), single built works (built environment), <single built works by specific type>, <single built works by function>, ceremonial structures, religious structures, religious buildings)

Classifications:
architecture (preferred)
Roman art

Creation Date: begun in 27 BCE, completely rebuilt 118/119-125/128

Creator Display:
unknown Ancient Roman, for the Emperor Hadrian (Roman emperor and patron, 76 CE-138 CE, ruled 117-138) [preferred,VP]
patron Hadrian, Emperor of Rome (Roman emperor, 76-138 CE) [500022200]
architect unknown Ancient Roman (Ancient Roman cultural designation) [500202843]

Locations:
Roma [7000874] Roma province (province), Lazio (region (administrative division)), Italy (nation), Europe (continent), World (facet) (Geographic)

Display Materials: constructed of stone, brick, concrete, and aggregate material; the drum is strengthened

Figure 94. Example of research result in the Getty CONA - I

Display Materials: constructed of stone, brick, concrete, and aggregate material; the drum is strengthened by huge brick arches and piers set above one another inside the walls
concrete [300010737]
.......(Materials Facet, Materials (hierarchy name), materials (matter), <materials by composition>, inorganic material)
stone (worked rock) [300011176]
.......(Materials Facet, Materials (hierarchy name), materials (matter), <materials by composition>, inorganic material, rock (inorganic material), <rock by form>)
brick (clay product) [300010463]
.......(Materials Facet, Materials (hierarchy name), materials (matter), <materials by composition>, inorganic material, clay, clay products)
aggregate [300014704]
.......(Materials Facet, Materials (hierarchy name), materials (matter), <materials by function>, additive (material))
rotundas (buildings) [300004842]
.......(Objects Facet, Built Environment (hierarchy name), Single Built Works (hierarchy name), single built works (built environment), <single built works by specific type>, <single built works by form>, <single built works by form: massing or shape>)

Dimensions: dome: 43 m (interior diameter and height) (141 feet); oculus: 8.9 m (diameter) (29 feet 2 inches)

Cultures:
Roman (preferred)

General Subject:
architecture (preferred)
religion and mythology

Specific Subjects:
worship [300056005]
.....(Associated Concepts Facet, Associated Concepts (hierarchy name), <religions and religious concepts>, religious concepts) (AAT)
Planetary Gods [1000079]
.....((Greek and Roman characters, Greek and Roman iconography, Religion, Mythology, and Legend)) (ICON)
Queen of Martyrs [1000080]
.....((Life of the Virgin Mary, New Testament narratives, Christian iconography, Religion, Mythology, and Legend)) (ICON)

Related Works:
depicted in .... Perspective Cutaway of the Pantheon, Viewed from the Northwest [700000165]
..........(preparatory drawing; Giovanni Antonio Dosio (Italian, 1533-after 1609); 1565/1568; Gabinetto disegni e stampe degli Uffizi (Florence, Firenze province, Tuscany, Italy))

List/Hierarchical Position:
..... Top of the CONA hierarchy
.... Built Works

Sources and Contributors:

Pantheon	.... [VP]
	.... CDWA online (1995-)
	.... Robertson, Handbook of Greek and Roman Architecture (1929)
Santa Maria ad Martyres	.... [VP]
	.... CDWA online (1995-)
	.... Robertson, Handbook of Greek and Roman Architecture (1929)
Santa Maria Rotunda	.... [VP]
	.... CDWA online (1995-)
	.... Robertson, Handbook of Greek and Roman Architecture (1929)

Subject: .... [VP]

Note:
English..... [VP]
..... Boethius and Ward-Perkins, Etruscan and Roman Architecture (1970)
..... CDWA online (1995-)
..... Sear, Roman Architecture (1983) 170 ff.

Figure 95. Example of research result in the Getty CONA - III

7.8.3 SWEET

Among the ontologies developed and published, we find the NASA (National Aeronautics and Space Administration) SWEET (Semantic Web for Earth and Environmental Terminology) ontology, intended for the representation of vocabulary in Earth and Environmental Sciences (EES) (Raskin, Pan, 2005). This ontology is modelled in the OWL language, and includes a very wide variety of concepts about both the studied object and the methodological aspects (research, analysis, measurements, residuals, etc.); these could be effectively transposed to other application fields among which is architectural heritage. Besides some general concepts about research and monitoring features, some entities could be used for describing certain characteristics of the monument itself (such as materials) or the context around it (which is actually an Earth and environmental issue). SWEET has become the *de facto* standard for data management in EES (Di Giuseppe, 2014).

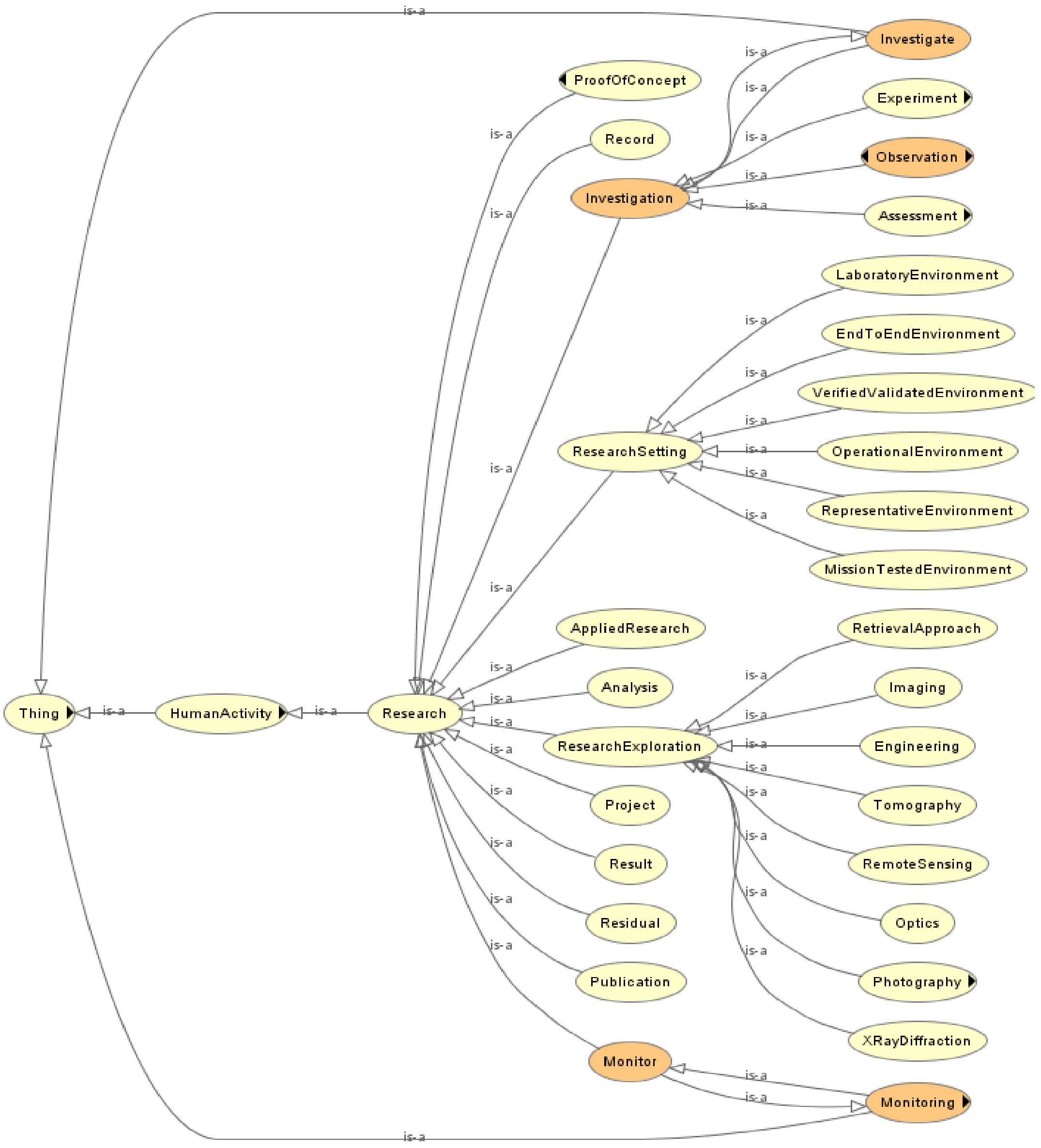


Figure 96. OWL Schema of the entity 'Research' in the SWEET ontology. Visualization in the OSS 'Protegé'[190].

[190] http://protege.stanford.edu/

7.8.4 Geonames

GeoNames[191] is a geographical database available and accessible through various web services, under a Creative Commons attribution license.

The GeoNames database contains over 10,000,000 geographical names corresponding to over 7,500,000 unique features. All features are categorized into one of nine feature classes and further subcategorized into one of 645 feature codes. Beyond names of places in various languages (Figure 98), data stored include latitude, longitude, elevation, population, administrative subdivision and postal codes. All coordinates use the World Geodetic System 1984 (WGS84) (Figure 97). Those data are accessible free of charge through a number of Web services and a daily database export.

The core of GeoNames database is provided by official public sources, the quality of which may vary. Through a wiki interface, users are invited to manually edit and improve the database by adding or correcting names, move existing features, add new features, etc.

Each GeoNames feature is represented as a web resource identified by a stable URI. This URI provides access, through content negotiation, either to the HTML wiki page, or to a RDF description of the feature, using elements of the GeoNames ontology[192]. This ontology describes the GeoNames features properties using the Web Ontology Language, the feature classes and codes being described in the SKOS language. Through Wikipedia articles URL linked in the RDF descriptions, GeoNames data are linked to DBpedia data and other RDF Linked Data in the Semantic Web.[193]

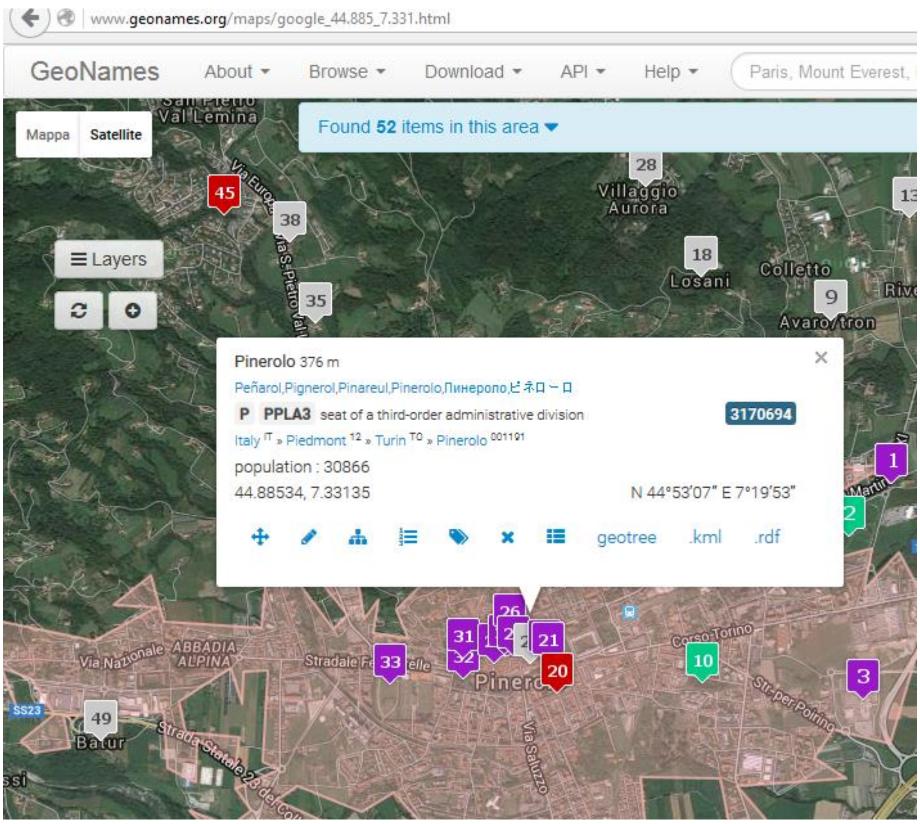


Figure 97. Example of results of a research inquiry on geonames.org – I: the map.

---

191 http://www.geonames.org/

192 http://www.geonames.org/ontology/documentation.html

193 The italic text cites https://en.wikipedia.org/wiki/GeoNames

3170694 Pinerolo

Feature | Hierarchy | History | Tags | **Alternate names**

code	lang	alternate name	p.	s.	h.	c.	action
							+
ja	Japanese	ピネローロ					+ ✎ ✖
es	Spanish	Peñarol					+ ✎ ✖
fr	French	Pignerol					+ ✎ ✖
pms	Piemontese	Pinareul					+ ✎ ✖
de	German	Pinerolo					+ ✎ ✖
en	English	Pinerolo					+ ✎ ✖
eo	Esperanto	Pinerolo					+ ✎ ✖
fi	Finnish	Pinerolo					+ ✎ ✖
it	Italian	Pinerolo					+ ✎ ✖
nap	Neapolitan	Pinerolo					+ ✎ ✖
nl	Dutch	Pinerolo					+ ✎ ✖
no	Norwegian	Pinerolo					+ ✎ ✖
pl	Polish	Pinerolo					+ ✎ ✖
pt	Portuguese	Pinerolo					+ ✎ ✖

Figure 98. Example of results of a research inquiry on geonames.org – II: the list of names.

7.9 OGC answers to the Semantic Web proposals

The world of geographic knowledge is working towards the exploitation of the newly available technologies. Some research has used Semantic Web technologies and languages for performing geospatial analysis (Zhao et al., 2008). Furthermore, some connected issues such as the voluntary mapping projects products can be integrated in a wider framework using ontologies (Du et al., 2011). Researchers in the geographic field are analysing the new offered opportunities (Laurini, Servigne, 2011). Finally, more extended projects aim to bring the geospatial data to the Web of Data. One of these is the GeoKnow project[194], which is funded by Europe in order to develop tools and features at the service of the web that exploit geographical information.

The same OGC, having acknowledged the potential of the Semantic Web technologies and open data, works to interface the existing standards with the new paradigm. From this effort, further standards and guidelines have been published. Some of them are GeoSPARQL (7.9.1) and the recently published Geospatial ontologies (7.9.2).

[194] http://geoknow.eu/Welcome.html

7.9.1 GeoSPARQL – a geographic query language for RDF data

The OGC GeoSPARQL standard supports representing and querying geospatial data on the Semantic Web. GeoSPARQL defines a vocabulary for representing geospatial data in RDF, and it defines an extension to the SPARQL query language for processing geospatial data. In addition, GeoSPARQL is designed to accommodate systems based on qualitative spatial reasoning and systems based on quantitative spatial computations.[195]

The GeoSPARQL standard follows a modular design; it comprises several different components:

- *a core component defines top-level RDFS/OWL classes for spatial objects;*
- *a topology vocabulary component defines RDF properties for asserting and querying topological relations between spatial objects;*
- *a geometry component defines RDFS data types for serializing geometry data, geometry-related RDF properties, and non-topological spatial query functions for geometry objects;*
- *a geometry topology component defines topological query functions;*
- *an RDFS entailment component defines a mechanism for matching implicit RDF triples that are derived based on RDF and RDFS semantics;*
- *a query rewrite component defines rules for transforming a simple triple pattern that tests a topological relation between two features into an equivalent query involving concrete geometries and topological query functions.*[196]

GeoSPARQL does not define a comprehensive vocabulary for representing spatial information. It instead defines a core set of classes, properties and datatypes that can be used to construct query patterns. Many useful extensions to this vocabulary are possible, and we intend for the GIS community to develop additional vocabulary for describing spatial information. Other standards groups are actively working to develop such a vocabulary (e.g. ISO/TC 211). [197]

7.9.2 OWS-10 Geospatial Ontologies

The core geospatial ontologies have been result of eight years of Research and Development at Image Matters in the domain of geospatial semantic. Image Matters decided to release these ontologies for the broad community seeking to facilitate semantic interoperability between systems using geospatial information. The intent is a bootstrap to the Geospatial Semantic Web, by providing an initial candidate geospatial ontology for standardization. We think a joint effort between OGC and W3C would be the right venue to bring this work to maturation and standardization due to their expertise both in geospatial domain and semantic web respectively.[198]

For each category, an IRI, a label, a description, restrictions and relations are provided, in order to foster their use in specifying geospatial characteristics in the semantic web. In the defined categories (Figure 99) it is possible to find again some of the concepts previously presented in the requirement description of the thesis subject.

[195] The italic text cites http://www.opengeospatial.org/standards/geosparql
[196] The italic text cites (OGC, 2012c)
[197] The italic text cites (OGC, 2011)
[198] The italic text cites http://ows10.usersmarts.com/ows10/ontologies/

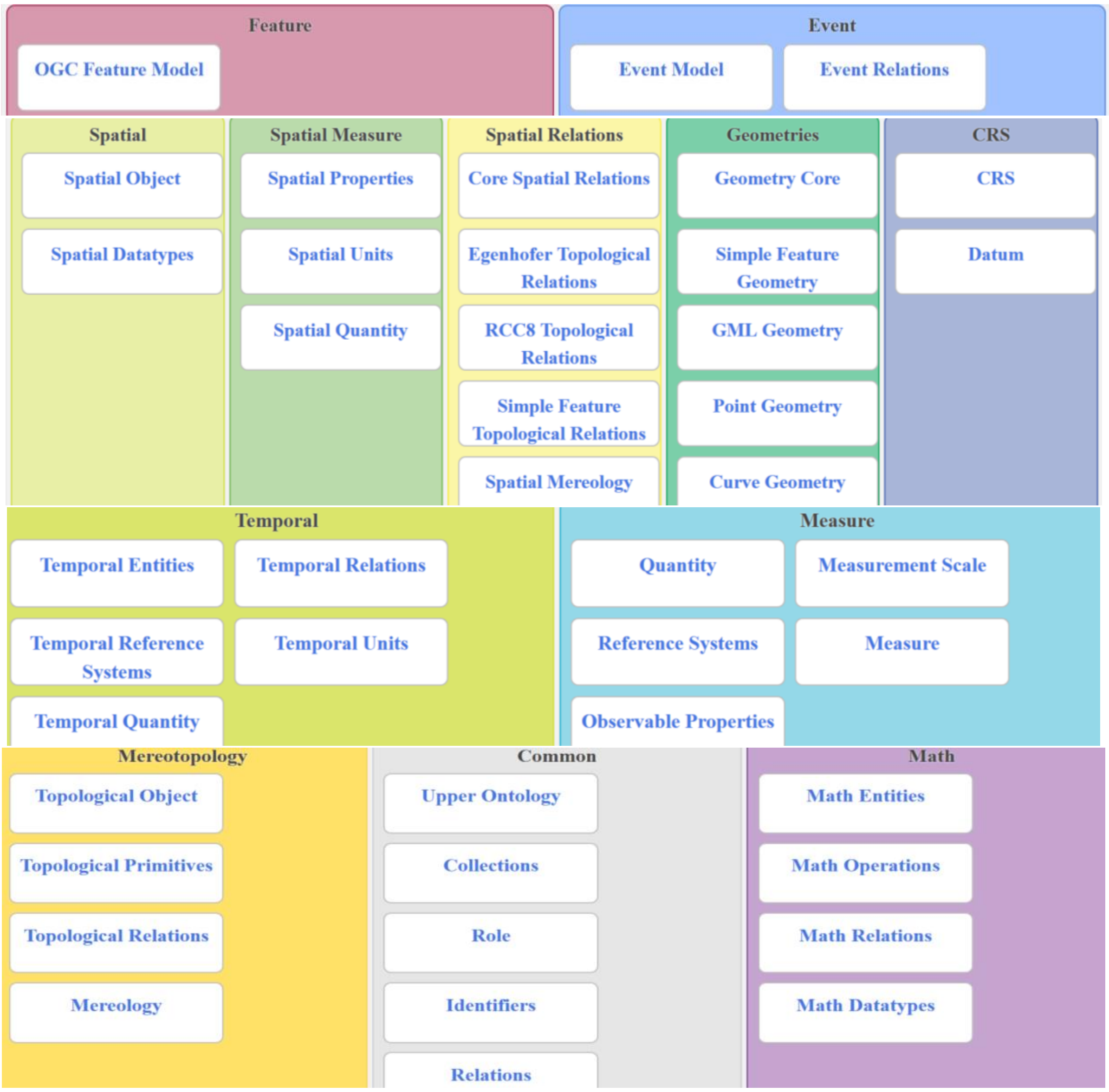
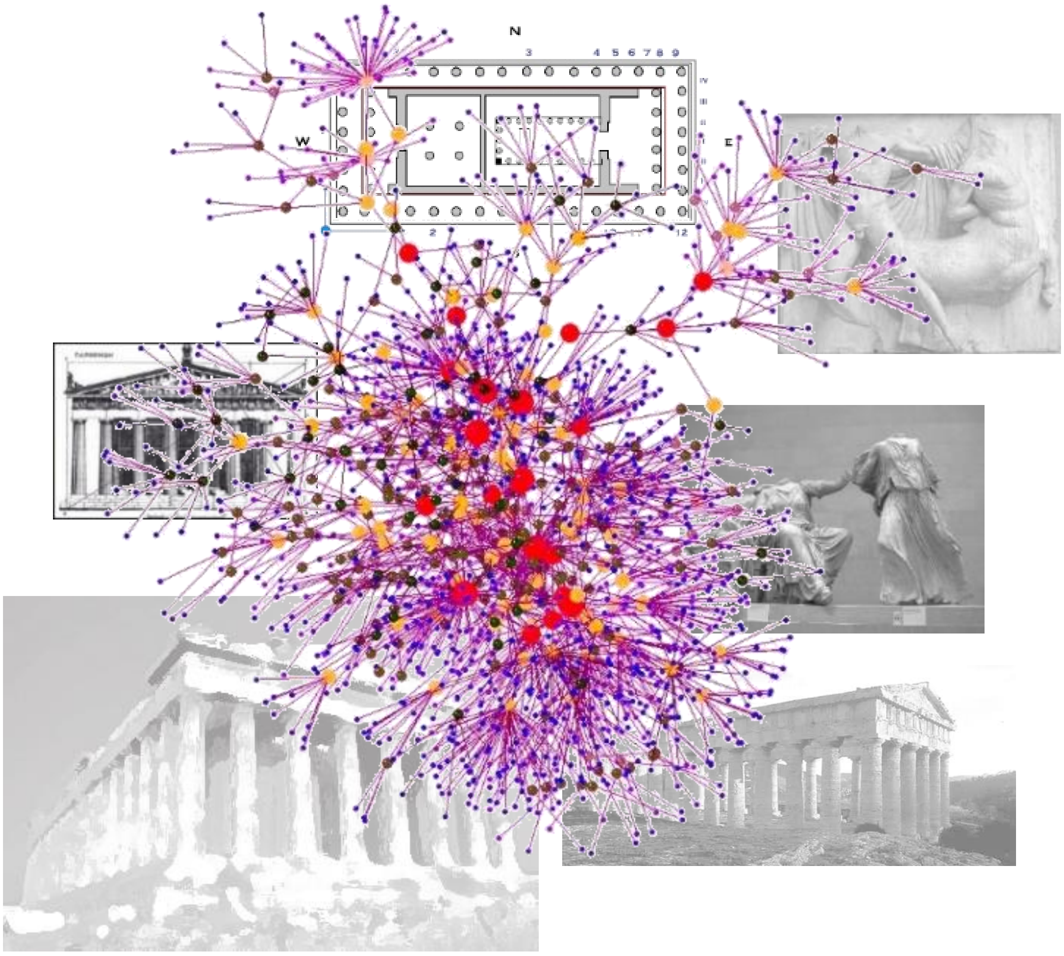


Figure 99. Geospatial ontology themes as in the web page[199].

[199] http://ows10.usersmarts.com/ows10/ontologies/

SECTION 3 – THE APPLICATION OF INTEROPERABILITY TECHNOLOGIES TO ARCHITECTURAL HERITAGE



(Adapted from http://static1.squarespace.com/static/55cc4c4be4b0dc6eee3b55ab/55cc4d58e4b07f2c67fbb18e /55cc4d5ae4b07f2c67fbb2a1/1327071057000/ontology.jpg?format=original)

8 A DATA MODEL FOR ARCHITECTURAL HERITAGE MANAGEMENT AND REPRESENTATION

Many structures and infrastructures have been published over the years and are available for the aims shared by the Semantic Web, the smart communities, and the standardization institutions and organisations, which can be summarized as: semantic interoperability (for database interoperability and information retrieval) and automated reasoning.

However, in the field of architectural heritage, an inclusive standard reference that could exhaustively represent these complex and multifaceted objects considering a high level of detail analysis, including the spatial documentation, is missing.

Two main fields meet in architectural heritage: cultural heritage and metric spatial data management, from cartography and regional or landscape analysis to the detailed 3D model. Other fields can participate in defining characteristics of architectural heritage, such as anthropology or economy; these will not be considered in detail, but some useful information, interesting to such disciplines, can be included, and the final models can be used with or linked to other specific representations.

Cultural heritage thematic aspects can be effectively represented using CIDOC-CRM structures. Instead, the spatial 3D models regarding architectures, which have great potentials of analysis and documentation, should be specified more to be exploited for all their documentation and analysis power.

For this reason, OGC CityGML was considered, since its structures already permit the representation of building features. It was extended to permit on the one hand, a more detailed and specific classification of the building components, and on the other hand, to be representative of the aspects that are important to the cultural heritage point of view. These include a number of multi-disciplinary issues and complexity elements. Moreover, the time connotation and the source documentation are introduced as important data for the documentation. This strategy was chosen, also because some implemented structures exist to realize the extension and to permit the verification of the results through visualization and analysis. The ADE mechanisms of CityGML are an affirmed technique for the realization of an effective model extension. Some software tools exist or are beginning to be developed for realizing and visualizing the system.

8.1 The CityGML Cultural Heritage Application Domain Extension

In this thesis, an extension of CityGML is proposed, in order to include the characteristics of surface complexity that are typical of architecture. Moreover, some attention is devoted to the traceability of the stored information, in order to include in the data the elements useful to technicians for interpreting the stored information and evaluating the degree of fuzziness of the data.

The CityGML extension is called the Cultural Heritage Application Domain Extension (CHADE). It is realized for the building module of CityGML (as summarized in Figure 100), but it could be similarly applied to the other CityGML modules. The **Errore. L'origine riferimento non è stata t rovata.** contains the full UML model. It is then analysed in detail in the following Section 8.1.1. The extension has been developed and will be tested on the building module; once its validity is proved, its concepts and classes can also be applied to the other CityGML modules.

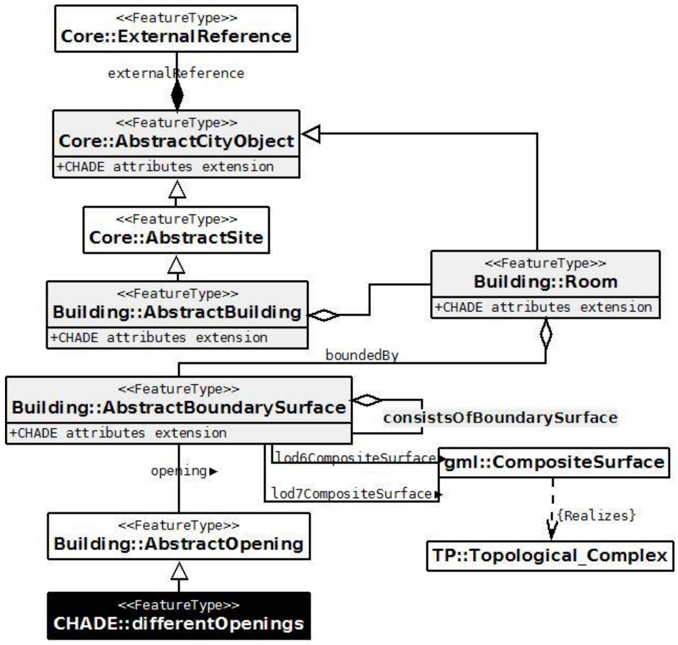


Figure 100. Synthesis of the CityGML CHADE. In white are the CityGML classes, in grey (black for the whole class) are the CHADE extensions and the inserted relations.

8.1.1 The CHADE components: research of granularity, flexibility and traceability

From the general to the particular, the first problem was to include some attributes useful for the identification of the monument and some related information (if a CH declaration exists, it should be determined what are the related documents, who are the owners, and what is the preservation authority). Some of these have been borrowed from previous investigations (Costamagna, Spanò, 2012), and extend the Core class ‘AbstractCityObject’ (Figure 101).

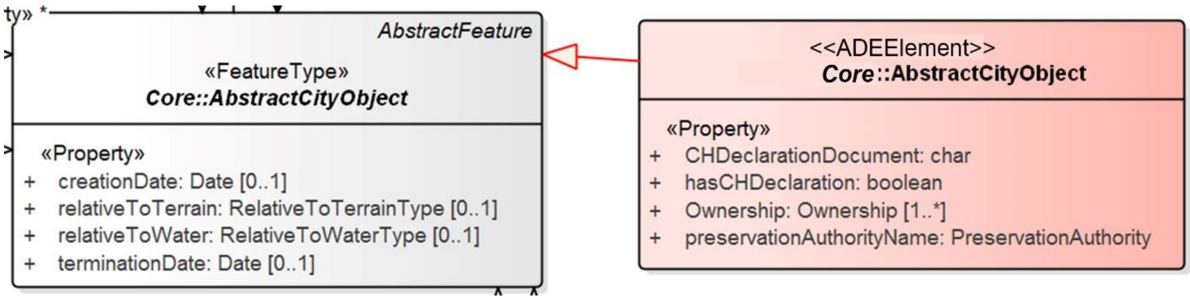


Figure 101. In grey, the CityGML class Core:AbstractCityObject; in red, the attributes that extend the class in the CHADE. The stereotype of the class should be ‘ADEElement’ following the instructions of the official procedures for processing ADEs.

The less-defined classes belong to the core module of CityGML, which is the parent of all the city objects, which inherit in this way its attributes and characteristics. A relevant class in the core model is the ‘ExternalReference’ class, already in CityGML, which permits the relation of the model with further databases managing data concerning the same object. For example, considering the management of the Versailles royal residence, the reference can be realized to the instance of the CONA vocabulary of the Getty Institute, which describes it.

The employed model structure permits the inclusion of complex attributes in the form of DataType, composed by a series of further attributes. This possibility is important for storing data about for example, the owners (Figure 102)

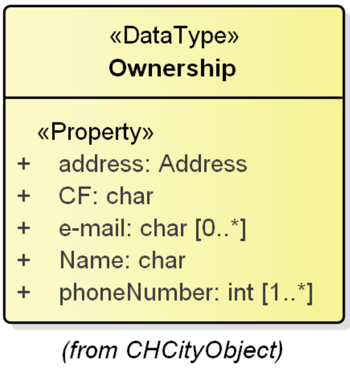


Figure 102. Example of data type introduced in the CHADE. The attribute 'Ownership' of the extended class 'AbstractCityObject' (in red in Figure 101) refers to this class.

The second issue is the extension of the attribute list for the 'AbstractBuilding' class. In particular, its function and its denomination are considered. Both of these values are complex when regarding a historical item, since both can change over time, and must be archived as a reference for studies and as an element for understanding the history of the building. Therefore, a DataType is included for both. The BLDG_Function data type includes at first the function name (at present, in English; specific future consultations with historians could further evaluate considering different languages in order to avoid losing the nuances of meaning). The reference to the URI of the Getty Institute vocabulary AAT (Art and Architecture Thesaurus) (7.8.2.3.1) follows, which includes the terms linked to the building's function as subclasses of *'single built works by function'*.

The last two attributes are present almost everywhere in the detailed added data types, because they are of fundamental importance for historical data connotation. The '*time'* attribute is defined as a time object defined in the same GML general schema. It could also be defined as a TM_Object (time object) as stated in ISO TC211 ISO 19108:2006 Temporal Schema, but some incompatibilities among some ISO TC211 definitions and GML requirements persist[200]. Nevertheless, both schemas have issues for detailing the time considered, as a date, as a period, with different degree of fuzziness and with the possibility to establish a sort of topology for temporal data, in a temporal reference system. This is of obvious importance for managing historical data.

The second attribute is '*Source'*, which is detailed, in turn, in a data type, including metadata, reference to the source, codes for its identification and retrieval, and the same attribute *'time'*.

Similarly, the attributes of the class '*Room*' are extended, adding '*RoomClass'*, '*RoomFunction'* and '*RoomUsage'*, all with reference to the Getty AAT vocabulary URI. The '*RoomUsage'*, which can change over time, is detailed in a dedicated data type.

Other references from AAT could be used for: '*Deterioration'*, from AAT '*components by general conext'*; for '*Material'* from AAT '*materials'*; for '*BSFunction'*, from AAT '*components by specific context'* (generally from '*system components'; 'building divisions'; 'architectural elements'; 'furnishing components'; 'visual works components'*).

The potentially more interesting part of the model is the extension of the CityGML class '*AbstractBoundarySurface*'. In the original model, it has no attributes, and can be specialized as belonging to the main parts of the buildings (e.g. RoofSurface, CeilingSurface, WallSurface). The

[200] https://en.wikipedia.org/wiki/ Geography_Markup_Language

change of this class can permit the description of the parts of the buildings to follow with significant flexibility, which are stratified and articulated, and even small portions can have different meanings. Therefore, a recursive "*part-of*" relation is added from AbstractBoundarySurface to the same AbstractBoundarySurface. This permits the articulation of the surfaces in hierarchical, semantically well-defined, multiscale, and possibly mereo-topologically defined parts. Several attributes are added and defined following the already explained criteria. Between these, the "*LevelOfSpecialisation*" (*LOS*) attribute deserves an explication.

The 3D models are usually considered for their geometric accuracy, which mainly derives from the production methods and measurement systems. This characteristic is stored in GML models as LoD (Level of Detail) associated to each geometry. Even if it implies some consequence on the level of semantic definition, it is mainly linked to the possibilities of representation offered by the available data, and thus to the accuracy and data density. The Level of Specialisation, is inserted in order to manage the possibility to define parts and subparts that can be recognisable on the same model (with homogeneous accuracy and LoD) but need to be separately specified because of the different meaning they assume if considered as a whole or as a singular part.

The "Level Of Development" used by BIM could seem to be a similar concept[201]. However, in BIM the more defined parts are newly modelled parts in a way that is similar to the changing of the representation for increasing the representation scale in cartography: a building is a point, then a squared polygon, then an approximation of a more articulated polygon, then a more accurate and sharped polygon with an articulated roof and heights. This when supposing that the original data were the same and the lower-scale maps are generated from the generalization of the more accurate survey data. Instead, the LoD as 'Level of Ddetail has to do with the quality of the input data, which can be used to produce different representations. In the proposed level of specialisation (LOS), the representation of the object remains the same, but the definition of the segmented parts increases with the growing of the LOS (Figure 103). Different semantic values are related to the parts of different LOSs.

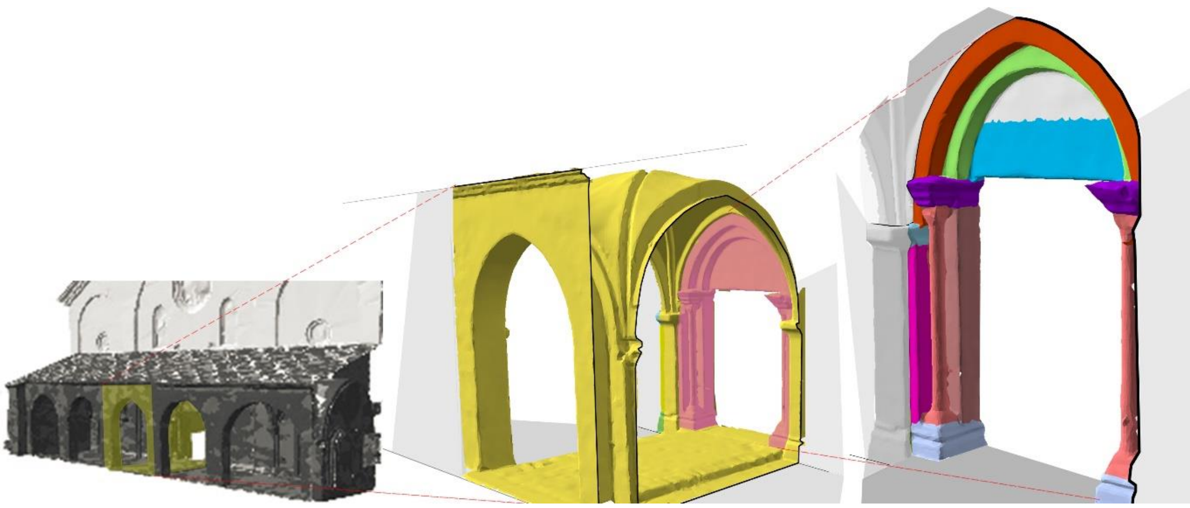


Figure 103. Examples of consecutive LOS specified on parts of a homogeneous-LoD 3D model. The colours represent the parts in which the object is divided.

A further extension of the class "AbstractBoundarySurface" regards the proper geometric levels of detail: two more LoDs are added, a *LoD5*, for approximately 1:200 - 1:100 scales and a *LoD6* for larger ones. The associated geometry class must be defined as a "*Geometric complex::GM_CompositeSurface*", since it is structured and hierarchical, in the same way that the

[201] http://bimforum.org/wp-content/uploads/2013/08/2013-LOD-Specification.pdf

boundary surfaces must also be semantically defined. Moreover, it has to be related to a "*Topological Complex::TP_Complex*", deriving from the "*Topology*" part of the standard ISO TC211 – ISO 19107:2003 Spatial Schema or the GML specification (they should be joined for the same issues described in the case of time objects). This last described part is complex when being used with current software and requires some additional efforts. Nevertheless, the inclusion of the topological relations as schematized in Figure 104 should be useful for correctly specifying the models.

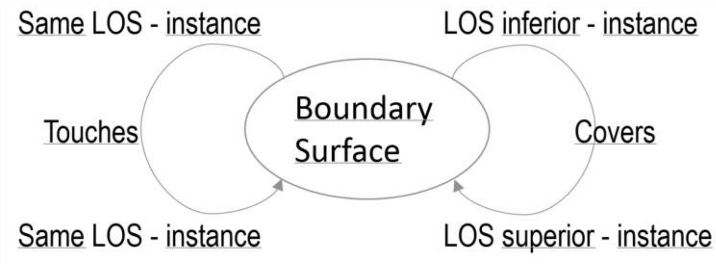


Figure 104. Schema of Egenhofer Topological Relations to be included in the model and verified. Some exceptions can exist (for example, a pillar can be considered only as one half as a component of a bay), these must be so analysed in order to confirm or not the validity of this model. The Egenhofer 2D topology has been used here even if the model is a 3D object because the single parts lie on the same surface (even if it is a 3D surface). The problem of how to treat different levels of details should be deepened (as stated in the previous sections). Moreover, 3D topology should refer to the 3D topology relations cited in Section 1.3.1.

Some considerations about the mereological relations that exist between higher LOS Boundary Surface (BS) and lower LOS Boundary surface according to classifications given in the section dealing with mereological concepts (Section 1.3.2) should be made. The nature of such relationships change following the semantics of the involved objects.

Fundamental for the relation between BS and BS is parthood relations, which also permit the representation of fiat parts (without 'bona fide', that is, defined, boundaries). This relation is characterised by density (there is always a smaller part) and transitivity (if A is part of B and B is part of C, A is also part of C). All these characteristics are essential to adequately represent the relation 'part of wall (e.g. a stratigraphic unit) – whole wall' (similar semantic), which is essential to express stratigraphy relations and the decay mapping.

Regarding examples about componenthood relations (between objects with different semantics), further classified by (Guizzardi, 2011), we can say that:

- A member – collective relation (Guizzardi, 2011) can include 'bricks – wall'.
- The component – functional complex relation (Guizzardi, 2011) exists between, for example, 'keystone – arch'.

In an extension of the model, this information should be included for improving the semantic enrichment and reasoning performances.

Concerning the two others relations defined by (Guizzardi, 2011):

- The relation subquantity – quantity (Guizzardi, 2011) (ex. 'sand – mortar') can be considered by ontologies representing materials or similar concepts. This could be of interest if for example a specific sand was used, for connecting the cave or particular manufacturing.
- The relation subcollective – collective (Guizzardi, 2011), regarding for example 'gemel windows – windows' does not affect the definition of Boundary Surface. It is not of interest

in this case for data modelling, since it is easily automatically inferred by considering hierarchical relations between semantic values.

"Containment" is the property that characterizes the relation between, for example, 'room – building' or 'central nave – church'. It is already contained in the CityGML model (and in the UML model, it is in fact represented as an aggregation) and could only be specified by adding possible semantic values to the objects in order to adapt it to architectural heritage representation.

OGC has processed some topological and mereotopological structure for helping to correctly store the data, but they are already oriented to linked open data formats, without regarding GML[202]. Moreover, the foreseen classes are oriented to the management of 2D objects.

In future work, an important part should be the link to external open data being structured as ontologies (in particular, the CIDOC-CRM and MONDIS could be useful for defining, for example, materials and deteriorations).

8.1.2 Implementation issues

The described model has been implemented using the method defined as the best practise by OGC (Van den Brink et al., 2012). UML schemas are modified that use stereotypes defined by a GML UML profile, so that their meaning can be understood by the machine and the performed transformation can be coherent and correct.

8.1.2.1 The extension in the UML model

For building the system, the proprietary commercial software '*Sparx Systems – Enterprise Architect'* is used. Contrary to the indications of using open source software for managing public (and open) data, it is also recommended for some official occasions (for example for the management of INSPIRE schemas). The software permits the importing of existing models (in this case, the CityGML building module and some general schemas such as GML are obviously used; also ISO 19108 for temporal objects and ISO 19107 for spatial issues could be considered, as well as the whole INSPIRE data model). The classes, selected and imported in the new extension model, maintain all their characteristics and relations with the other parts of the model they belong to. This is crucial to avoid creating an isolated new model, but permitting the insertion into a complex existing framework.

From this basis, new classes can be added, the attributes can be defined and new relations can be established.

The used types for defining the domains of each attribute should be compatible with GML, or the generated XSD will not be valid.

One particular data type that can be meaningful for interconnection reasons is '*anyURI'*: for vocabulary references (in particular, to the Getty Vocabulary terms), the 'anyURI' type was used. It is a string type that compiles with the lexical restrictions of a URI. Some attributes in GML are based on anyURI: `xlink:href`, `srsName`, `uom` and `frame` These are used to reference remote data. `srsName`, `uom` and `frame` reference remote dictionary entries and can be attached to objects or

[202] http://ows10.usersmarts.com/ows10/ontologies/

properties. A `xlink:href` attached to a GML property indicates that the value of the property is pointed to by the `href` attribute (Lake et al., 2004).

The so-formed model can then be exported in different formats, including XSD for being used as a GML application schema. Other interesting formats are OWL, ArcGIS workspace, and similar.

8.1.2.2 The editing of the generated XSD

Contrary to what it seems from this synthesis, the implementation steps are not so easy, or, better, they have to be controlled and corrected, since they require suitable settings and reference files describing in the specific software-understandable language how the transformation must be done. Waiting for this progress, possibly planned as future work, the resulting files have to be manually corrected by editing the XML text of the XSD file.

Three XML documents are generated: a comprehensive schema citygml_CH.xsd, which includes the others, that is, the CHCityObject.xsd and the CHBuildings.xsd (**Errore. L'origine riferimento non è stata trovata.**).

The problems encountered in the generation of the XSD from the UML schema built in *Enterprise Architect* are presented in the following.

The first problem was in the use of the stereotype '`ADEElement`'. It was initially used for the classes created as dependencies of the existing classes, with the same name, in order to add attributes to these last ones. But, notwithstanding the OGC best practice for modelling ADE in UML and the instructions of some tutorials available on the web, the software employed (*Enterprise Architect*) was not able to recognise it. The reason is probably that a specific CityGML UML profile is required to be used as a coding reference for performing the translation, but it is difficult to find it on the web (possibly it was licensed under some software house property). A GML UML profile is instead available, but the definition of `ADEElement` is excluded, and the simple editing of the XML by adding a stereotype similar to the others is not effective. It is probable that some more complex definition is necessary, since the class under that stereotype must be transformed for becoming part of a further existing class. This involves more than simply creating a new class.

For the development a stereotype '`feature type`' has therefore been used, but, obviously, the 'is-a' relationship is interpreted as a childhood relationship. In the XSD a substitutionGroup is added to the class definition element, meaning that it is the parent class (Figure 105)

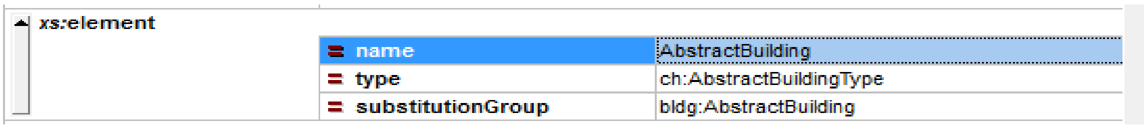


Figure 105. Example of CHADE class interpreted as a child of AbstractBuilding in the XSD.

It is necessary to change this incorrect interpretation making the system understand that the class only extends an existing class of another referred model (CityGML). For doing that, the XML text was manually edited using XML editors such as *ALTOVA XMLSpy* or the free and open source *XPad*, even if a simple text editor like *Notepad* could be used. In analogy with the XML definition of the attributes that extend some classes in existing ADEs, the definition of the class as a child was eliminated, and each new attribute was defined with the addition of a `substitutionGroup` defined as:

GenericApplicationPropertyOf[classe di riferimento]

(es. bldg:_GenericApplicationPropertyOfAbstractBuilding).

Some examples are reported in Figure 106.

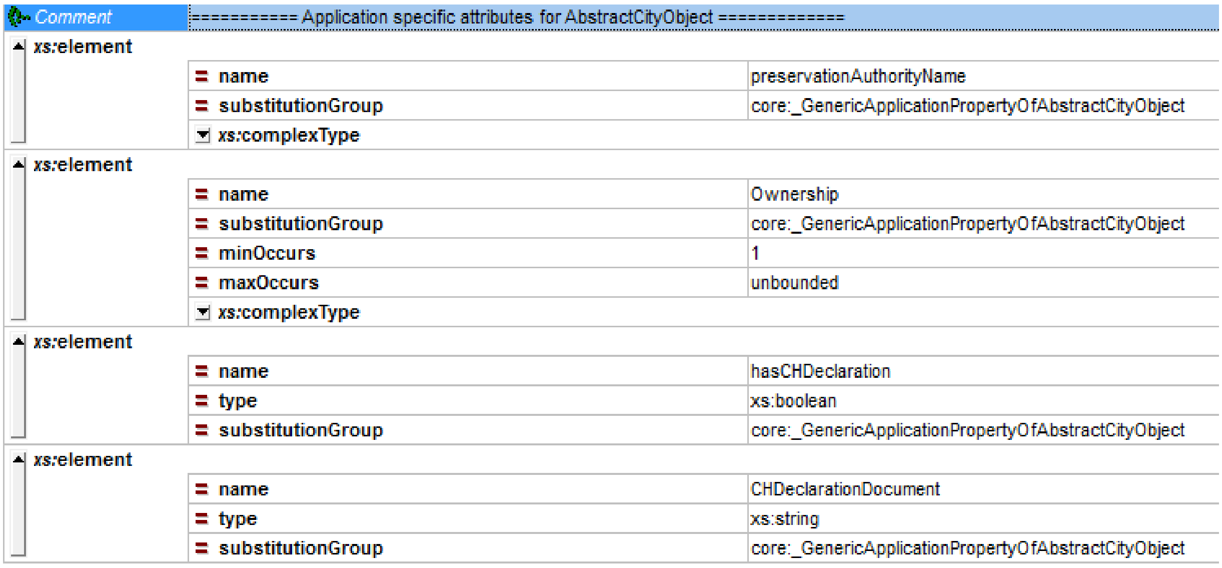


Figure 106. Examples of attributes that extend a CityGML class in the XSD.

The attributes can be defined in this form. However, some problems are given by the setting of "`minOccurs`", since the *XMLSpy* validation gives the error "Attribute '`minOccurs`' is not allowed in element `xs:element`". The use of a complex type should be defined (Figure 107), but it should be verified if the systems reading the XSD can correctly read it.

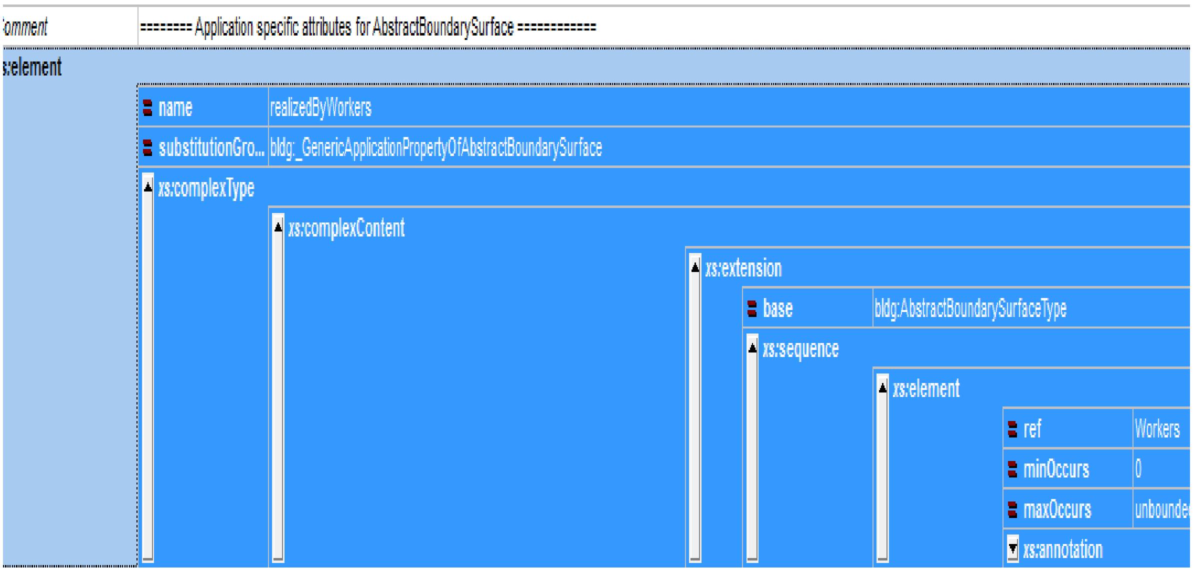


Figure 107. Example of a complex objet to define cardinalities in the attribute definition.

A further issue is the necessity to define the `XMLNS` (XML namespace) in the heading part of the document. It is defined as a local position 'D:DOTTORATO/CH_ADE'.

Furthermore, as already mentioned, some different standards have proven to be incompatible with GML and, consequently, with CityGML. Among the schemas to be harmonized are, for example, topology (GML topology.xsd must be used), time, and addresses.

Also, the reference to the INSPIRE addresses schema, which could also be used for being connected to lower-scale maps, was not found. For defining addresses, the CityGML reference to OASIS xAL (extensible Address Language)[203] was used (Figure 108). In future works, an external reference to Geonames should be added.

```
<?xml version="1.0" encoding="UTF-8"?>
<!--Modified by Ram Kumar (MSI) on 24 July 2002-->
- <xs:schema elementFormDefault="qualified" xmlns="urn:oasis:names:tc:ciq:xsdschema:xAL:2.0"
  xmlns:xs="http://www.w3.org/2001/XMLSchema" targetNamespace="urn:oasis:names:tc:ciq:xsdschema:xAL:2.0">
  - <xs:annotation>
      <xs:documentation>xAL: eXtensible Address Language This is an XML document type definition (DTD) for defining
        addresses. Original Date of Creation: 1 March 2001 Copyright(c) 2000, OASIS. All Rights Reserved
        [http://www.oasis-open.org] Contact: Customer Information Quality Technical Committee, OASIS
        http://www.oasis-open.org/committees/ciq VERSION: 2.0 [MAJOR RELEASE] Date of Creation: 01 May 2002
        Last Update: 24 July 2002 Previous Version: 1.3</xs:documentation>
    </xs:annotation>
  - <xs:annotation>
      <xs:documentation>Common Attributes:Type - If not documented then it means, possible values of Type not
        limited to: Official, Unique, Abbreviation, OldName, Synonym Code:Address element codes are used by groups
        like postal groups like ECCMA, ADIS, UN/PROLIST for postal services</xs:documentation>
    </xs:annotation>
```

Figure 108. Reference to OASIS xAL in the XML definition (the targetNamespace at the 4th line).

Some attention should be paid to the employed version of CityGML[204]. The problem is that not all the applications reading the CityGML files are sufficiently updated to read the more recent versions. In the realized test the version 2.0 of CityGML and, consequently, of cityGMLBase.xsd was used[205]. The reference must be the same in all the referenced schemas.

8.1.2.3 The XML validation

Once these aspects were edited and corrected, the XML was validated by the XML editors *ALTOVA XmlSpy* and the open source software *Xpad*.

This means that the XML syntax is correct and that all the reference to external XML schemas can be read[206].

A synthesis of the requirements for a well-formed XML document is given in http://www.w3schools.com/xml/xml_validator.asp:

An XML document with correct syntax is called "Well Formed".
The syntax rules are:

- *XML documents must have a root element;*
- *XML elements must have a closing tag;*
- *XML tags are case sensitive;*
- *XML elements must be properly nested.*[207]

[203] https://www.oasis-open.org/standards
[204] in http://www.citygmlwiki.org/index.php/ XML_namespaces_and_schemaLocations the available versions can be checked.
[205] http://grepcode.com/file/repo1.maven.org/maven2/org.citygml4j/citygml4j/2.1/schemas/CityGML/2.0.0/cityGMLBase.xsd
[206] https://www.w3.org/TR/xmlschema-1/
[207] The italic text cites https://www.w3.org/TR/xmlschema-1/

9 MANAGING A 3D MODEL IN CITYGML CHADE: PREPARING AND FILLING IN THE 3D DATA

Once the model is built and verified, the data about a monument, collected with various techniques, including the dense high-scale 3D models, must be subject to some preparation phases that permit their management in the CityGML CHADE.

The problems of the full workflow are often linked to the inability of the software to manage some functionalities and algorithms and/or formats at the same time. For this reason, several processing steps using specific software tools are needed.

In the following, a case study is considered for illustrating the phases of processing of the 3D models (assuming as a starting point an already processed and optimized 3D mesh) managed in the described system.

9.1 The case study

The considered case study is the Cistercian Staffarda Abbey (Figure 109), a monastery located near Saluzzo in northwest Italy that was founded in the XII century by Cistercian monks. The abbey complex grew larger between the XII-XIII centuries, with a gradual decline from that date onwards (Rotunno, 2011). In 1750, the Holy See declared the autonomous abbey's role to be over, and the complex was given to the Order of St. Maurice. Due to the modifications over the centuries, the monastery combines both Roman and Gothic architectural styles, and it includes the Santa Maria Church, a cloister, other monastic rooms (the dormitory, the refectory, etc.), the covered market, and the guestrooms (Beltramo, 2010). Much of the complex was built in red brick and red sandstone. The church has a nave and two aisles, and is a splendid example of the Romanesque-Gothic style.

The inside is austere, and the cross vault and the pillars (all different from one another) are decorated with alternating colours that which range from red to grey.

Externally, the overload of the roof led to the construction in early medieval sites of the flying buttresses that have improved but not solved the problem. Currently, the monastery represents one of the most important testimonies of medieval architecture in Piedmont. A number of studies were realized on the complex, and a wide documentation exists. Some of the analysis already uses geomatics techniques as a support for studies of surfaces (Donadio, Spanò, 2015).

A detailed survey was performed using integrated techniques (Bastonero et al., 2014), and a complete 3D model of a part of the church was then available in form of a textured mesh (Figure 110) (model processed by E. Donadio; Donadio, 2013).


.I

.II

Figure 109. The Cistercian Staffarda Abbey: (I) From the main road; (II) The church façade.

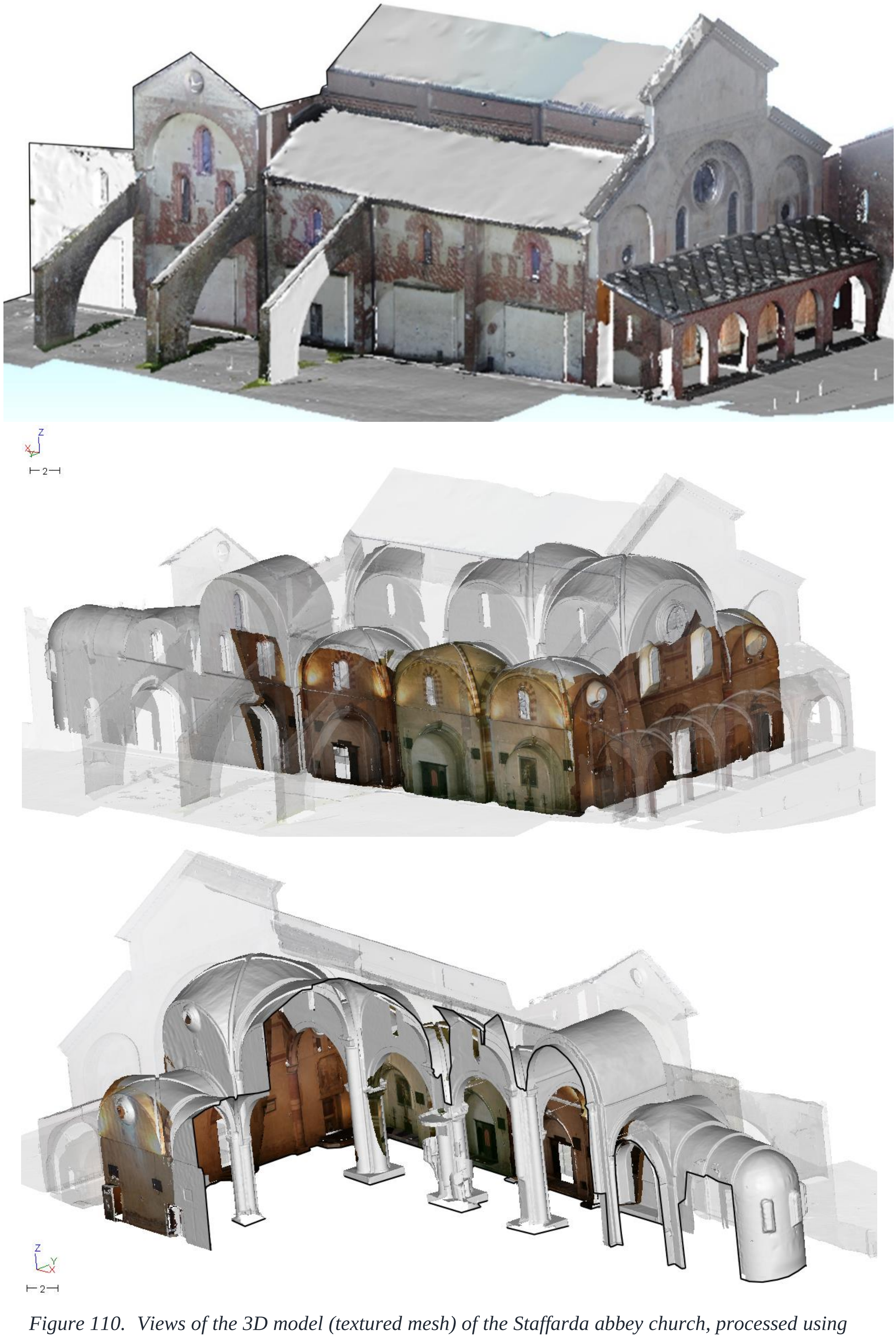


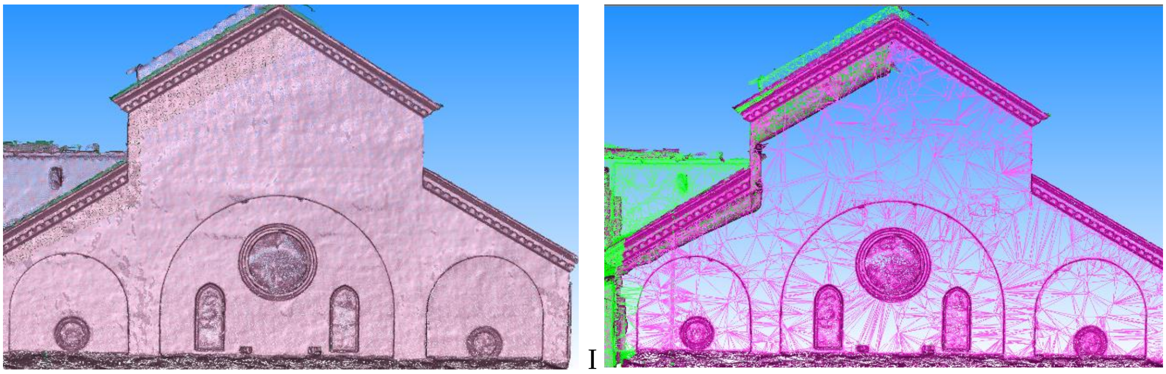
Figure 110. Views of the 3D model (textured mesh) of the Staffarda abbey church, processed using LIDAR acquisitions integrated with photogrammetric data acquired from UAV (Unmanned Aerial Vehicle) (Bastonero et al., 2014) (model processed by E. Donadio; Donadio, 2013).

9.2 Reduction of the computational intensiveness of the products

Even in 2D digital cartography, some simplification and generalization of objects are applied when downscaling for readability and visualisation reasons. For these processes, some generalisation rules are followed (McMaster, Shea, 1992).

The first editing phases to be performed on the models regard, on the one hand, the reduction of their points, considering the conservation of the original definition (Figure 111). This process can be performed using different algorithms, not always known in proprietary software. Consequently, this topic should be further analysed in order to establish the methods and the limits of this practice.

Some studies exist on the level of simplification that can be applied to models to be sufficiently complex for being expressive and understandable by the human eye on the one hand, but at the same time not too heavy to be managed on the web or in other computational environments (Apollonio et al., 2010).



I II

Figure 111. 3D models of the church façade before (I) and after (II) reduction. Both models view are in wireframe modality (which only shows the triangles edges).

9.3 Semantic Segmentation

On the other hand, the model must be segmented, so that every part must be isolated from the others for being suitably managed geometrically and semantically (Figure 112 - Figure 118); moreover, the temporal connotation must be considered in the segmentation, since it is essential for historical objects to map stratigraphy. A surface can be eventually repeated if it is part of more than one instance having different LOS as attributes. Also, this field can offer a quantity of techniques that should be analysed for identifying the most suitable one.

The following images show the 3D model (in form of a texturized mesh) that is progressively cut with the aim of segmenting the parts to be managed in the system as successive LOSs. The views in the following figures are screen shots from the software *Hexagon 3DReshaper*, which was used for processing the model.

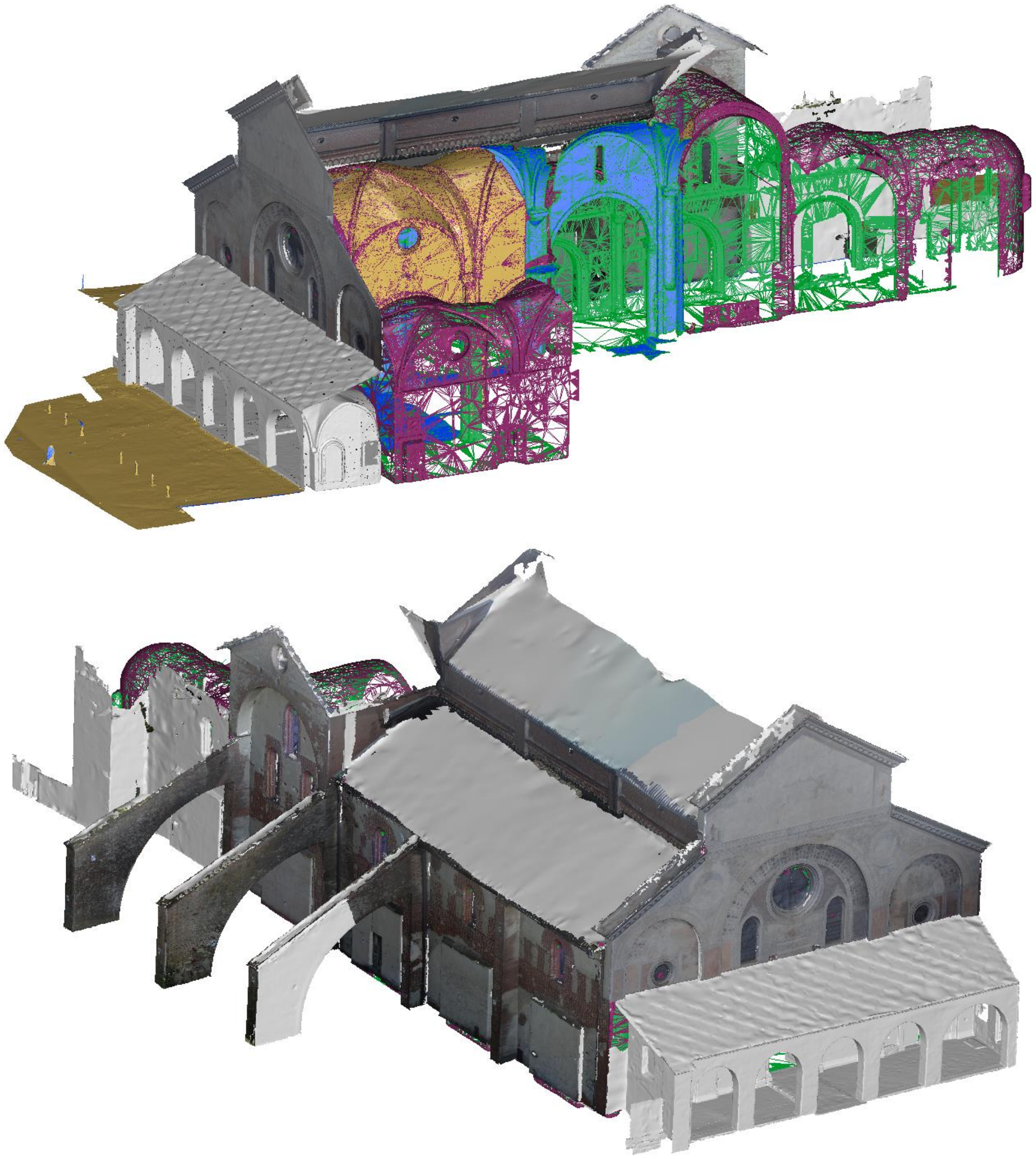


Figure 112. Full 3D model of the abbey (interior + exterior) (LOS1) in a rapresentation scale about 1:50. The geometries are visualized in texturized modality, filled mesh or simply wireframe modality (which only shows the triangles edges); anyway all the parts have similar nature. The different colours highlight some already made segmentation.

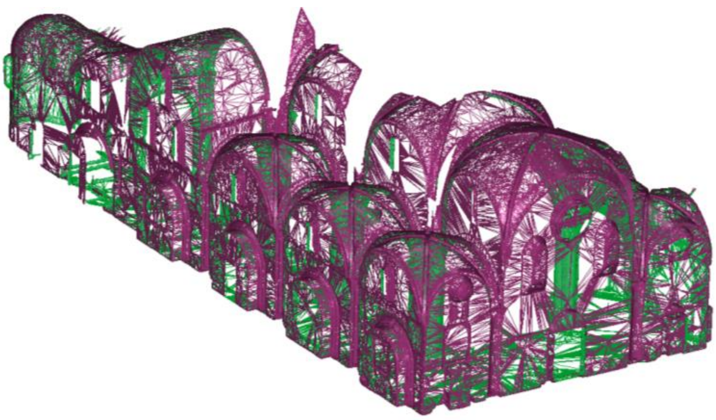


Figure 113. Interior of the church (LOS 2).

Figure 114. Only the narthex is visualized (LOS 2). Here, the model of the terrain is also visible and could be included in the CityGML model.

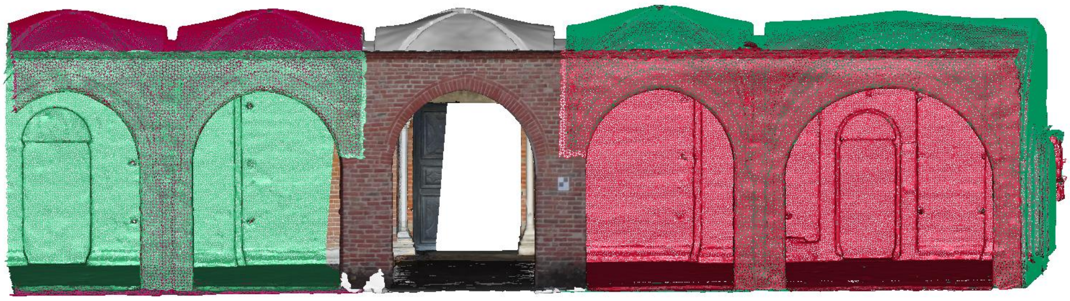


Figure 115. The narthex is segmented for considering the bay as a unit of the entire porch (LOS3).

Figure 116. The bay further segmented in its parts (LOS 3)

Figure 117. (I) The column is highlighted as it is segmented and represented in LOS 4; (II) The representation deriving from a further survey (realized with digital images processed in a structure-from- motion software) is integrated in the 3D model of the bay. It is realized for being segmented for LOS 5 - 6.

Figure 118. 3D model of one capital with segmented parts in different colours (LOS 6).

These phases can be performed in 3D model processing and editing software, such as *Hexagon 3D Reshaper* or similar applications. Some software can compute directions or the normal of the mesh faces. A classification of these values could be of interest for automatically defining constituent parts (Figure 119).

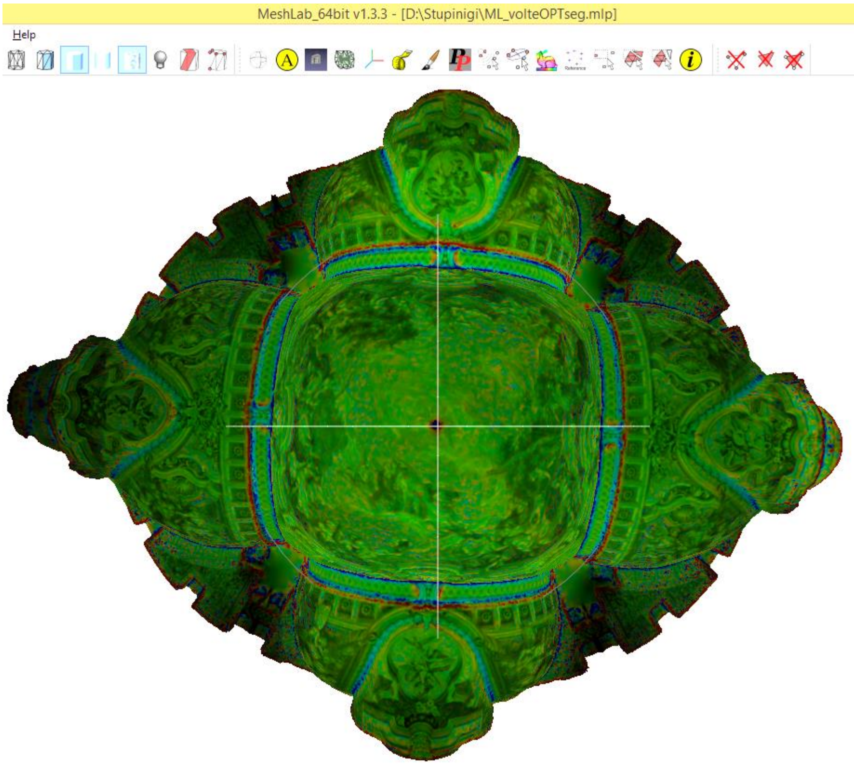


Figure 119. Texturized mesh of the vault of hall of honour of the Stupinigi royal residence, a hunting castle of Savoy real house near Turin (Italy). The colours map the discontinuities between surface portions direction. The software (in this case the open source software Meshlab was used) computed that by doing statistics on adjacent faces normal and directions.

Segmentation algorithms and procedures are being developed. (Hichri et al., 2013) present a summary of some of the available techniques used by algorithms for segmenting point clouds:

"One of these methods applied in the architectural field is based on color similarity and spatial proximities (Zhana et al., 2009): it uses an algorithm based on region growing in order to find the nearest neighbor of each seed point creating regions which will be merged and refined on the basis of colorimetrical and spatial relations.

Another method is based on shape detection (Ning et al, 2010): In a first step, an algorithm based on region growing and normal vectors is adopted to segment each planar region. Then, architectural components are extracted through an analysis of planar residuals.

There are also another method based on a distance measured between planar faces (Dorninger et al., 2007). This method is inspired from the 2.5D segmentation approach introduced by (Pottman et al., 1999) and it measures the distance in order to determine seed-clusters for which a region growing algorithm is performed. After that, an analysis of component connection is accomplished in the object space in order to merge similar seed-clusters.

Previous point cloud segmentation are limited to surfaces segmentation. In the field of cultural heritage, studies are almost not diffused and not very relevant. However, in the field of industry, many researches focused on this issue and presented interesting results (Golovinskiy, Funkhouser, 2009).[208]

Particular attention should be devoted to some methods directed to the recognition of specific parts of meshes through the concepts expressed by ontologies, in the form of topology and geometry of the parts (Attene et al., 2009). It would be interesting to develop similar processes for architectural models. The segmentation could therefore be realized using ontologies. Future work should be addressed to this topic.

9.4 Technical issues for editing of models

The processed model must be exported in COLLADA format for the following transformation. This XML-based open exchange format[209] is not managed by some proprietary software, so the model must be exported in 3D model format (such as OBJ or PLY) and reimported in further software able to perform the exportation. For example, the open source software *MeshLab* can accomplish that.

The problem is that it has difficulties in managing high coordinate values, so that the whole model must be translated near the origin for this passage.

The exported COLLADA files can then be reimported in GIS management systems, such as *ArcGIS*, as multipatch shapefiles (ESRI, 2008). Here, they have to be moved again to their original position in georeferenced coordinates.

9.5 The translation to CityGML CHADE format

Since the managed surfaces are complex, being composed, although already reduced, by miles of triangles (stored in the form of rings composing a multiple composite surface), it is obviously not possible to manually store the singular points, but they have to pass through a series of phases that permit their export in a GML format. Few software applications are at present available for performing this passage.

The most widespread application is the proprietary software *FME* (licenced by *Safe Software*). *FME* is probably the most advanced software for transforming various formats of 3D models into CityGML structures. The parts of the model can be mapped to the CityGML entities based on implemented libraries, and a CityGML dataset is automatically generated[210]. The limitation is that customized extensions cannot be used to map the objects, since it is not possible to modify the software source.

A second tool is available to transform the 3D model into the CityGML format. That is the use of the *ESRI ArcGIS* '*Data Interoperability'* extension. This permits the exportation of the segments of the 3D model in a generic CityGML format. In this case, the algorithms of *FME* are also used, being integrated in the *ArcGIS toolbox*.

[208] The italic text cites (Hichri et al., 2013)
[209] https://www.khronos.org/collada/
[210] http://www.safe.com/how-it-works/

While *FME* can read a number of 3D model formats, *ESRI ArcGIS* requires the ESRI shapefile format for enabling editing. Therefore a first exportation from the COLLADA format, which can be imported in *ArcGIS* but minimally modified, and ESRI Multipatch format must be realized. Then, some attributes can be added, if necessary, and some edits can be applied. For example, since the reference system of the model was changed, in the *ArcGIS* environment it can be set and the model can be moved to its correct coordinates again. Subsequently, the parts of the model can be exported in a CityGML format.

The result is the inclusion of the geometry and the attributes of the single parts of models in files structured as CityGML, semantically classified as generic 'CityObjects' and indexed with a created ID. The GML files (readable as XML structured text) have to be manually modified for including the CHADE in the description schemas references, and to correctly define the semantics of each part.

9.5.1 The need for the manual editing of XML

The GML files produced by *ArcGIS* include the reference to all possible CityGML modules, probably because the possibility to set this kind of requirement is reduced (including all the possibility is a good method for not to get wrong). The reference to the new extension must be added (with the `xmlns` suffix '`ch'`); it is a local path (D:/DOTTORATO/CH_ADE).

The first part of the document refers to the envelope of the georeferred model, with the reference system expressed as an EPSG code[211].

The rest of the document presents the objects in the form of attributes and geometries (the origin was a relational database), classified as generic 'CityObjects'.

The hierarchy must be set and the correct labels must be applied following the CityGML file format. Moreover, all the textual attributes must be manually filled in.

Another important issue is to add suitable identifiers, in order to realize some connections, for which, for example, `Xlink` is required.

The IDs must uniquely identify each object; for this reason, it is important to define some rules to follow in their composition. In the case study, they are structured as follows:

Dataset name [= object + creation year + operator]_Element [= (record=REC, attribute=AT, geometric attribute=GE) + LoD (Dn°) + LOS (Sn°) + progressive number 000000]

ex. Staff15Polito_RECD4S1000001

That is: Staffarda + 2015 +Politecnico di Torino _ Record + LoD 4 + LOS 1 + number 000001.

An example of URI could be similar to the following:

http://www.polito.it/id/WallSurface/Staff15Polito_RECD4S1000001

[211] www.epsg.org

9.5.2 The validation of the GML file

The following rules are similar to the ones cited for XSD validation, and the GML models are also tested. Another proof of their validity is the possibility of being read by visualisation software.

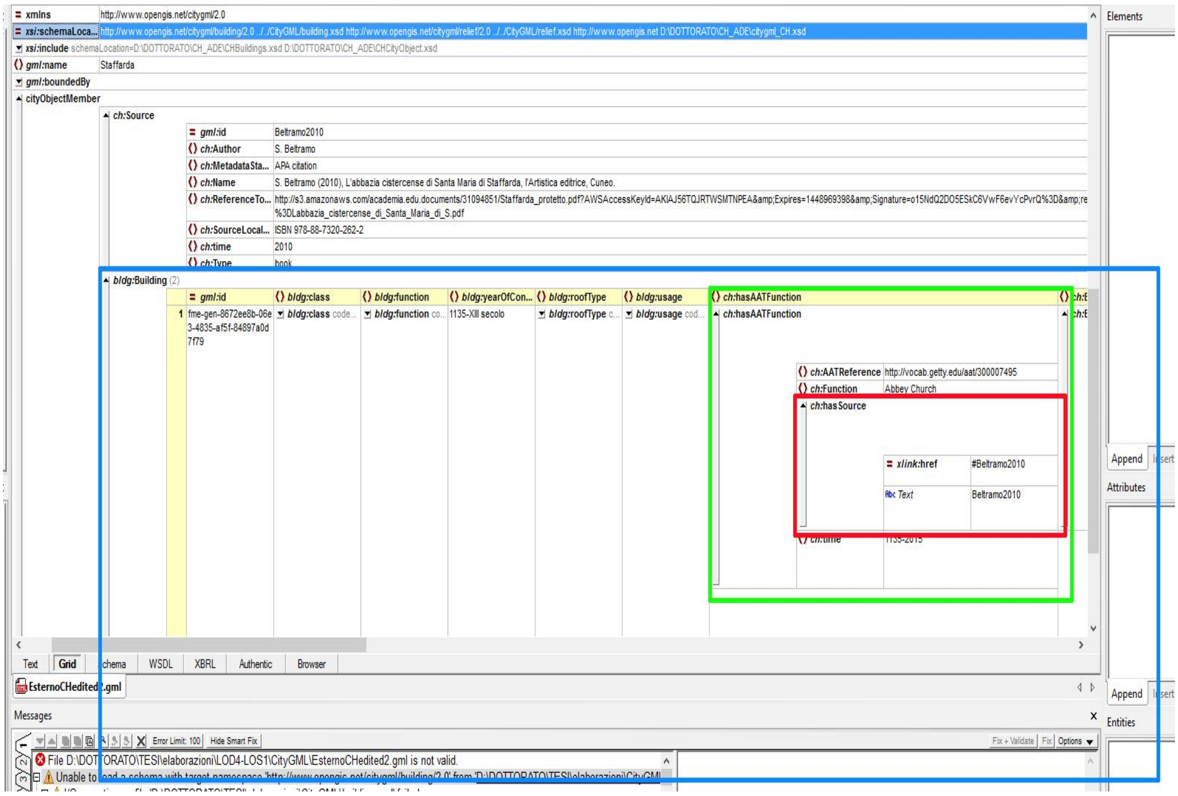


Figure 120. Final GML documents in which the included objects are underlined: The blue object contains the green object, which is linked to the red one.

9.5.3 A note about texture

CityGML permits the storage of a texture applied to 3D models, using the module ‘Appearance’. In the meantime, structure-from-motion systems permit the generation of models with very high-resolution textures (even for very complex objects), which can unequivocally provide richness for the meaning of the model and serve as a fundamental element for architectural heritage representations. The problem, not addressed in this thesis, could be the format in which the texture is produced and that must be inserted in CityGML. As a reflection for future studies, an example of texture automatically produced and exported by a common SfM software (*Agisoft PhotoScan*) is represented in Figure 121. It is clear that transformations and further considerations or issues must be taken in account in order to permit the use of such a texture in the model.

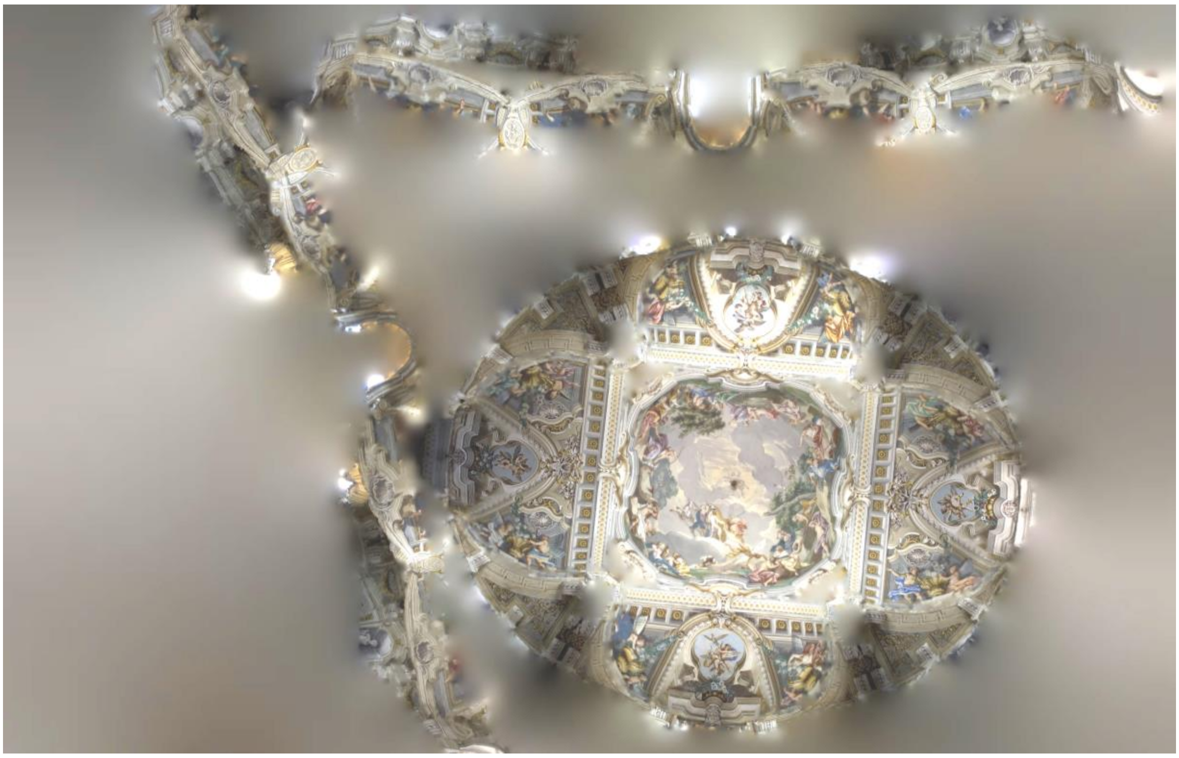


Figure 121. Texture automatically exported by an SfM software model of the Stupinigi hall of honour.

10 RESULTS: SURFING THE ARCHIVE. VISUALISATION AND ANALYSIS IN A GRAPHICAL INTERFACE

At this point, the archive, in the form of a GML file, can be shared through the web and read by several applications or interfaces for being consulted and analysed. For this goal, several options are available.

The first one is to query the document in its XML format, using `XQuery` sentences in XML processors.

Another way is to query the system through the graphical interface for extracting the contained information in a way similar to more traditional GIS interfaces.

A further option could be the development of further programs and applications that could read the archive and realize more queries or processing on it. This could also be developed as a part of the Semantic Web, in which the archive should be published for being read and queried by all the Semantic Web users.

The second option is here reported, with some examples about the possible information retrieved through the open source application *FZK Viewer*, which is one of the most advanced alternatives, especially in the form of open source software. It is a free CityGML and IFC viewer application developed by the Institute for Applied Computer Science, KIT – University of the State of Baden-Württemberg, and National Laboratory of the Helmholtz Association (formerly named Forschungszentrum Karlsruhe). It implements the schemas of some versions of IFC and CityGML, including some official CityGML ADEs (such as the Noise ADE). It has an open structure that permits being customized by adding other CityGML schemas to be used. In this case, for example, the CHADE schema was added in the directory of the reference files of the software for the described data to be interpreted. This is unequivocally an advantage of the open structure of the software.

A list of similar existing software tools can be reviewed in http://www.citygml.org/?id=1538.

The chosen tool offers several functionalities that permit the user to surf the model, and, at the same time, to test the goodness of the realized GML file employing the proposed CityGML CHADE.

In Figure 122 the archive is visualized in the FZK Viewer, which permits the reading of both the geometry, at different levels of detail, and the associated attributes. The relations in which the object is involved can be queried by the platform, and the results can be easily visualized and read, in both the geometric side and the thematic attributes (Figure 123). The geometry of the model permits the collection of measurements directly on the objects. In Figure 124 some of them are visible: the automatic measurement of areas or distances.

Also, non-geometrical objects can be read. When the relation is not a geometrical object, as is the case, for example, of the data Source, only the attributes are listed as a result of the relation (Figure 125). This is an advantage if one is planning to have large bibliographic reference lists and large datasets: a church can be cited in a book (or document) about one topic, and in the same book, other similar objects can be cited; if all of them are related, some information retrieval for specific interests would become easier. Furthermore, if the source is linked to the collection where it is preserved, or to information for buying or borrowing it, the efforts of the researcher would become easier.

Following the same mechanism, and activating the link, the connection to other databases and archives, such as the Getty Vocabularies or Geonames, could be activated for accessing the resource.

Moreover, the tool permits the automatic generation of some statistics and synthetic tables related to the managed objects (Figure 126 - Figure 127).

Some thematic visualizations can be applied, for example, for visualizing the year of construction (Figure 129). A similar query could be used to verify if, in a similar epoch, for example, narthexes were added in many churches (perhaps having common type, or function, shape, or geographical position, etc.), possibly enlightening rules, or events, or tendencies in all similar objects. This becomes possible if large similar datasets will be realized and shared in the Semantic Web using this extended ontology as reference for data modelling.

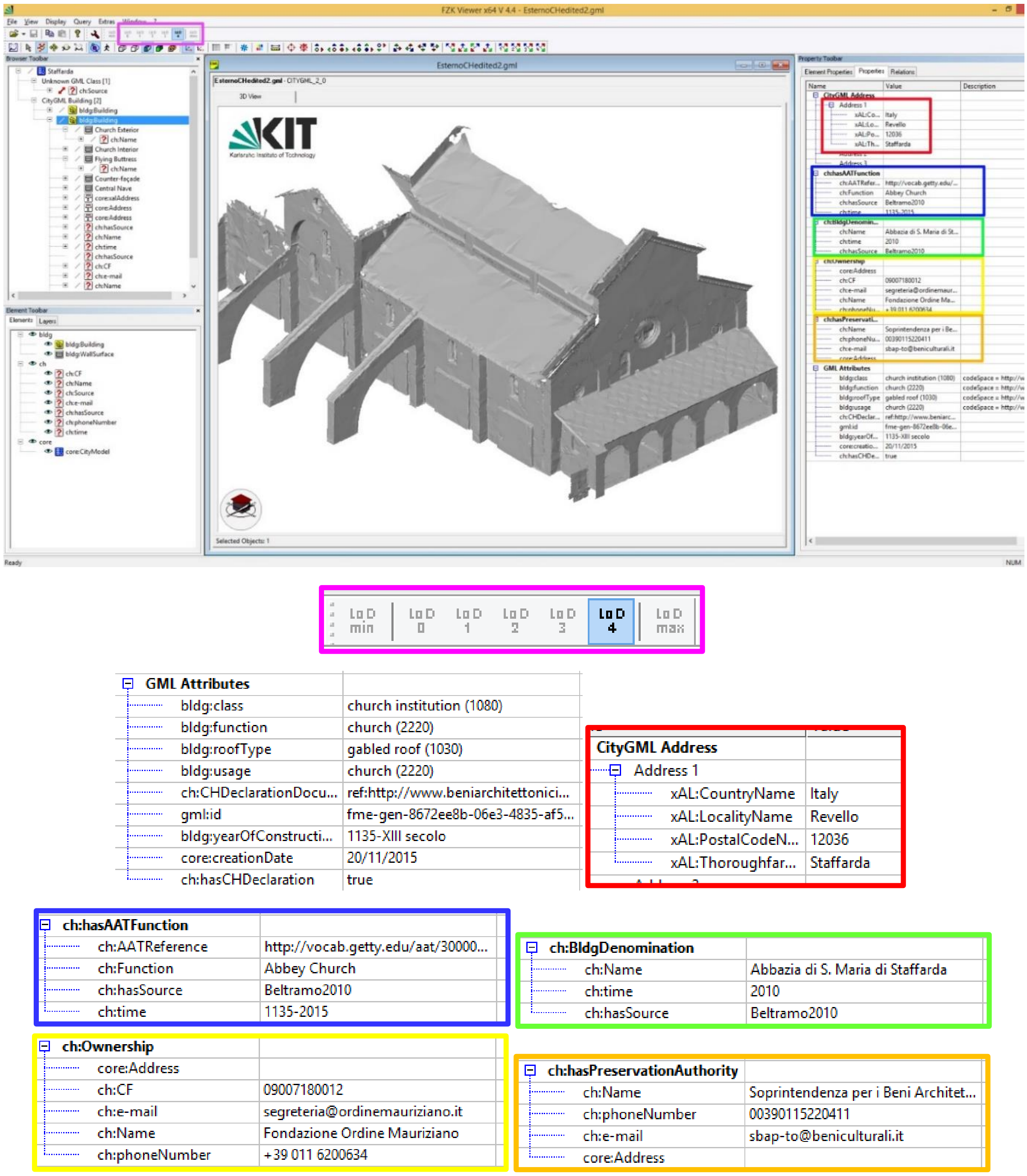


Figure 122. Visualization of the archive in the FZK Viewer: in the centre, the geometry of the object is visualized; on the left side, it is possible to read the hierarchy of the contained objects, and on the right, the properties can be read. In different colours, the attributes that are in turn objects themselves or data types (and are therefore composed by a set of attributes) are highlighted. The level of detail to be visualized can be chosen, since the data are multi-scale (in purple in the figure).

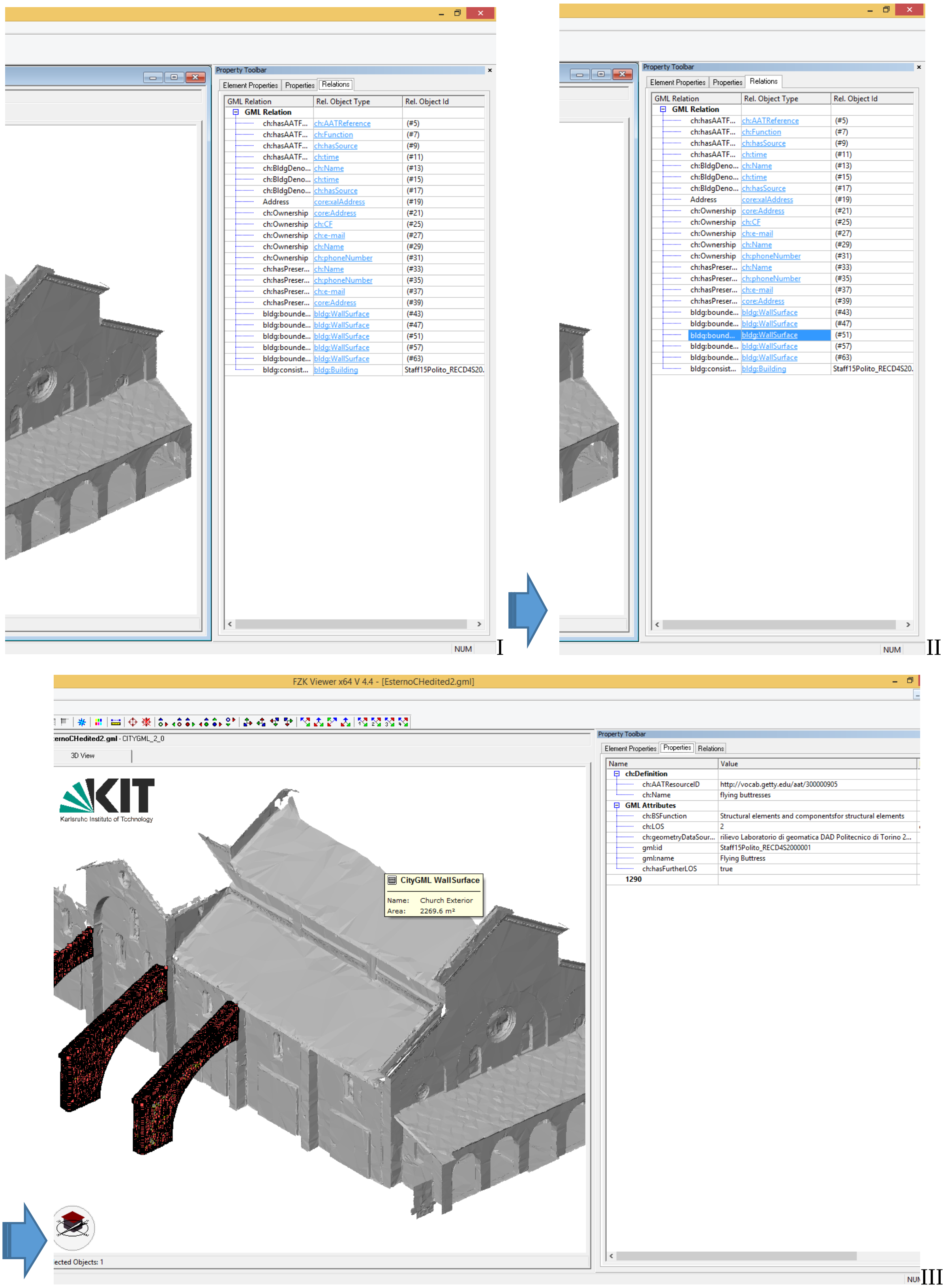
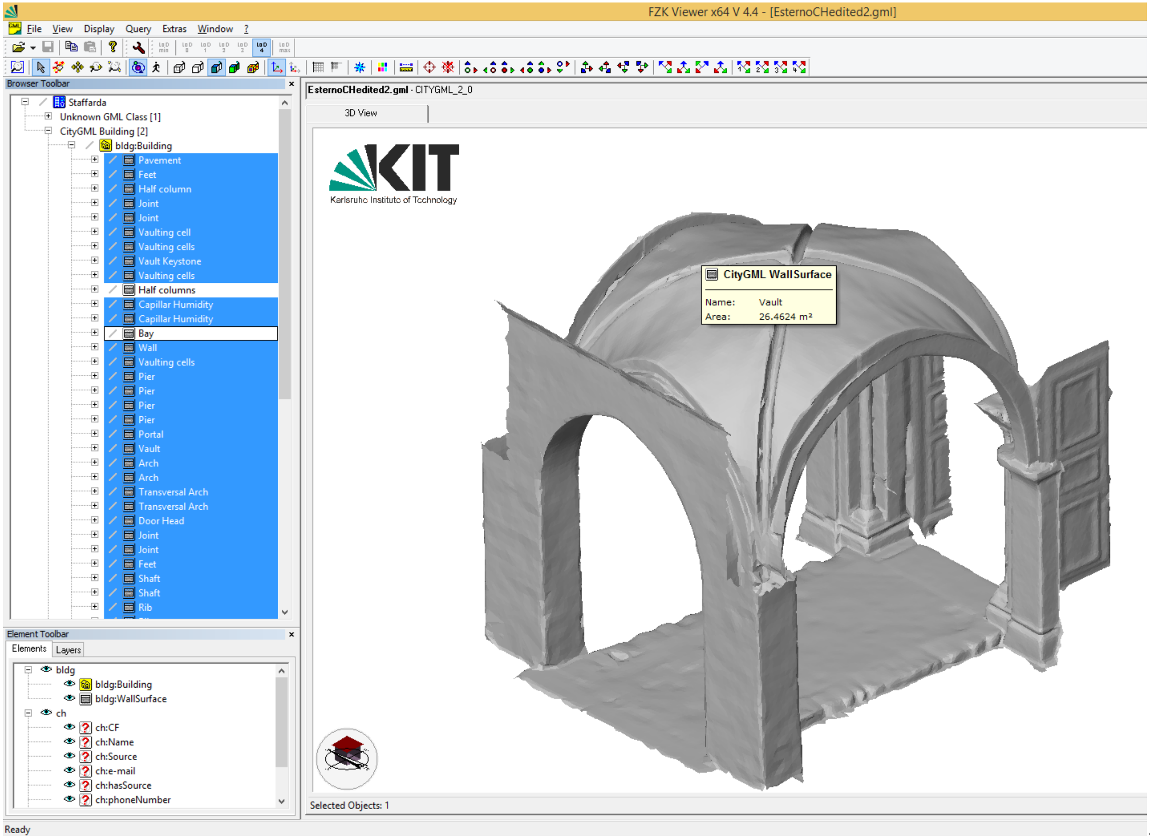
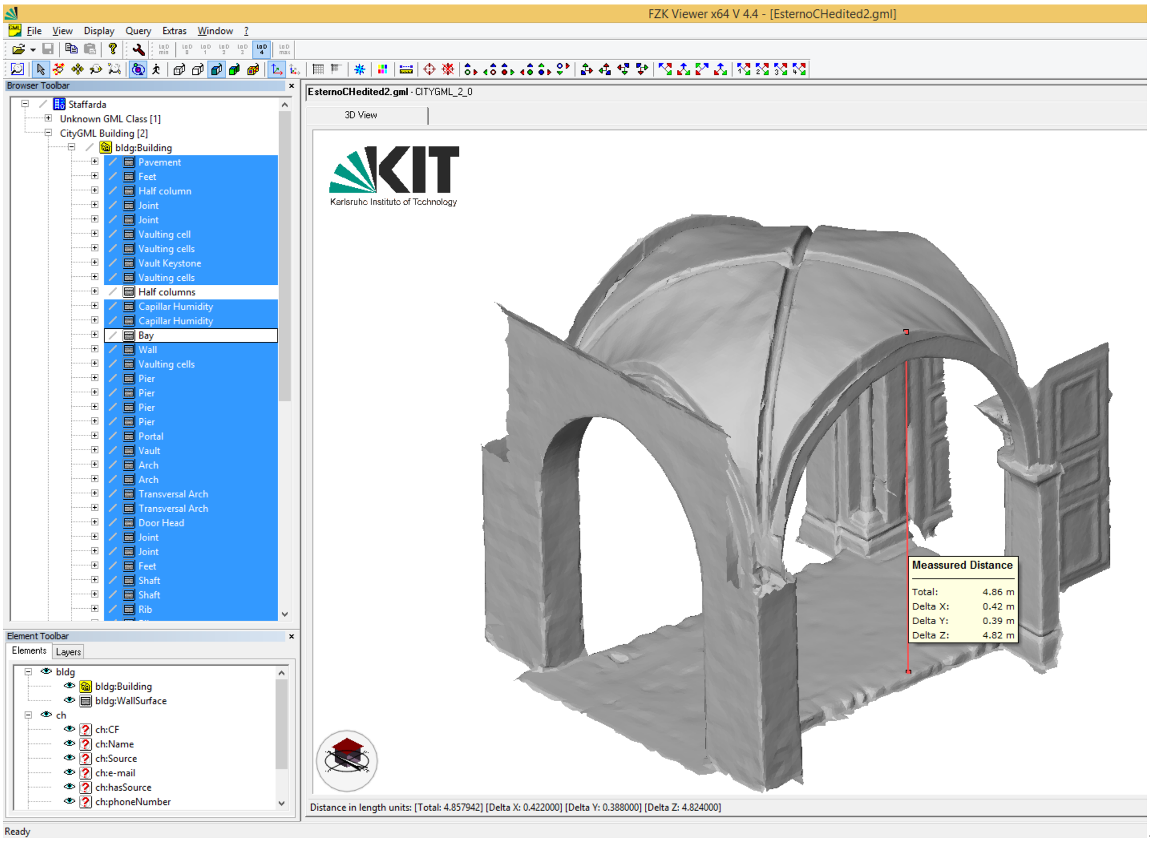


Figure 123. (I) The window is visible, where relations to further objects are contained. If one of them is selected (II), the object properties (geometry and thematic attributes) are visualized. The result of the link is given by the relation of the whole object to one of its parts (the flying buttress). They are selected in the representation, and the attributes are listed in the right part (III).



.I



.II

Figure 124. Examples of measurement possibilities are shown: when the mouse cursor passes on one surface, the total area of that surface is shown (I); linear distances can be manually measured (II).

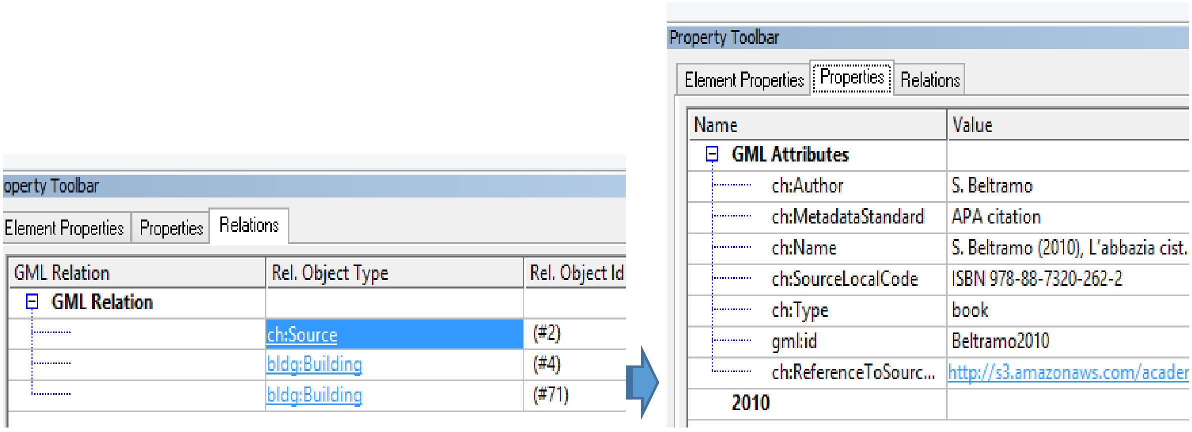


Figure 125. Example of result of a relation to an object containing only textual attributes.

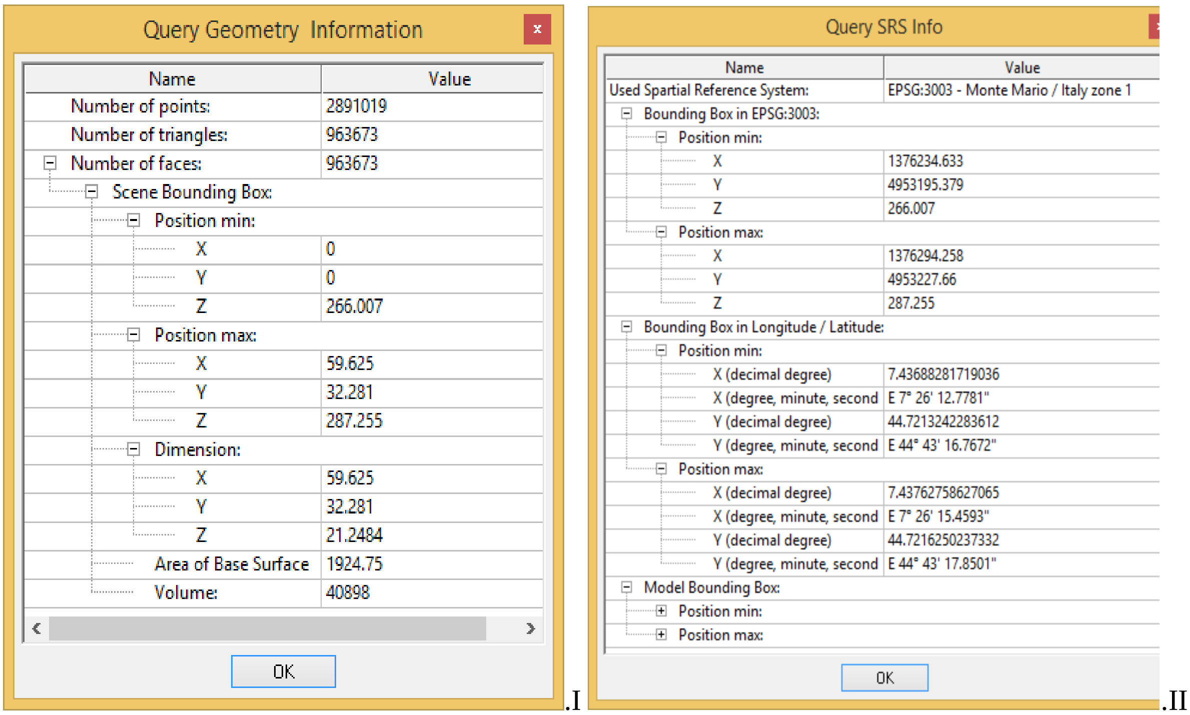


Figure 126. Some automatic statistics that can be computed by the system: geometry statistics (I) and SRS statistics (II).

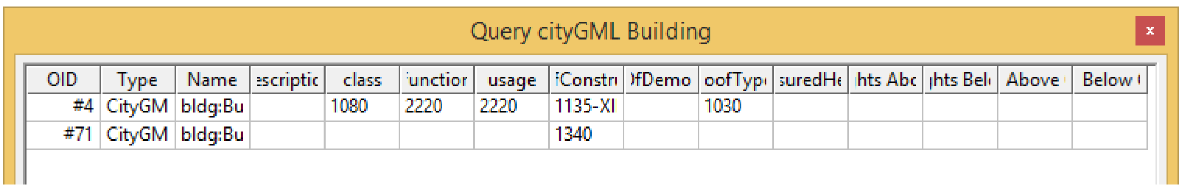


Figure 127. Query about classified CityGML entities

Query Entities

OID	Type	Name	Description	Geometry Types
#1	CityGML Modell	Staffarda		
#2	Unknown GML Class	ch:Source		
#4	CityGML Building	bldg:Building		
#5	GML Data Type	ch:AATReference		
#7	GML Data Type	ch:Function		
#9	Unknown GML Class	ch:hasSource		
#11	GML Data Type	ch:time		
#13	Unknown GML Class	ch:Name		
#15	Unknown GML Class	ch:time		
#17	Unknown GML Class	ch:hasSource		
#19	CityGML Address	core:xalAddress		
#21	CityGML Address	core:Address		
#22	GML Data Type	xAL:AddressDetails		
#25	Unknown GML Class	ch:CF		
#27	Unknown GML Class	ch:e-mail		
#29	Unknown GML Class	ch:Name		
#31	GML Data Type	ch:phoneNumber		
#33	Unknown GML Class	ch:Name		
#35	Unknown GML Class	ch:phoneNumber		
#37	Unknown GML Class	ch:e-mail		
#39	CityGML Address	core:Address		
#40	GML Data Type	xAL:AddressDetails		
#43	CityGML WallSurface	Church Exterior		CompositeSurface | Face
#44	Unknown GML Class	ch:Name		
#47	CityGML WallSurface	Church Interior		
#48	Unknown GML Class	ch:Name		
#51	CityGML WallSurface	Flying Buttress		MultiSurface | Face
#52	GML Data Type	ch:AATResourceID		
#54	Unknown GML Class	ch:Name		
#57	CityGML WallSurface	Counter-façade		MultiSurface | Face
#58	Unknown GML Class	ch:Name		
#60	GML Data Type	ch:alternativeSource		
#63	CityGML WallSurface	Central Nave		
#64	Unknown GML Class	ch:Name		
#66	GML Data Type	ch:AATResourceID		

OK

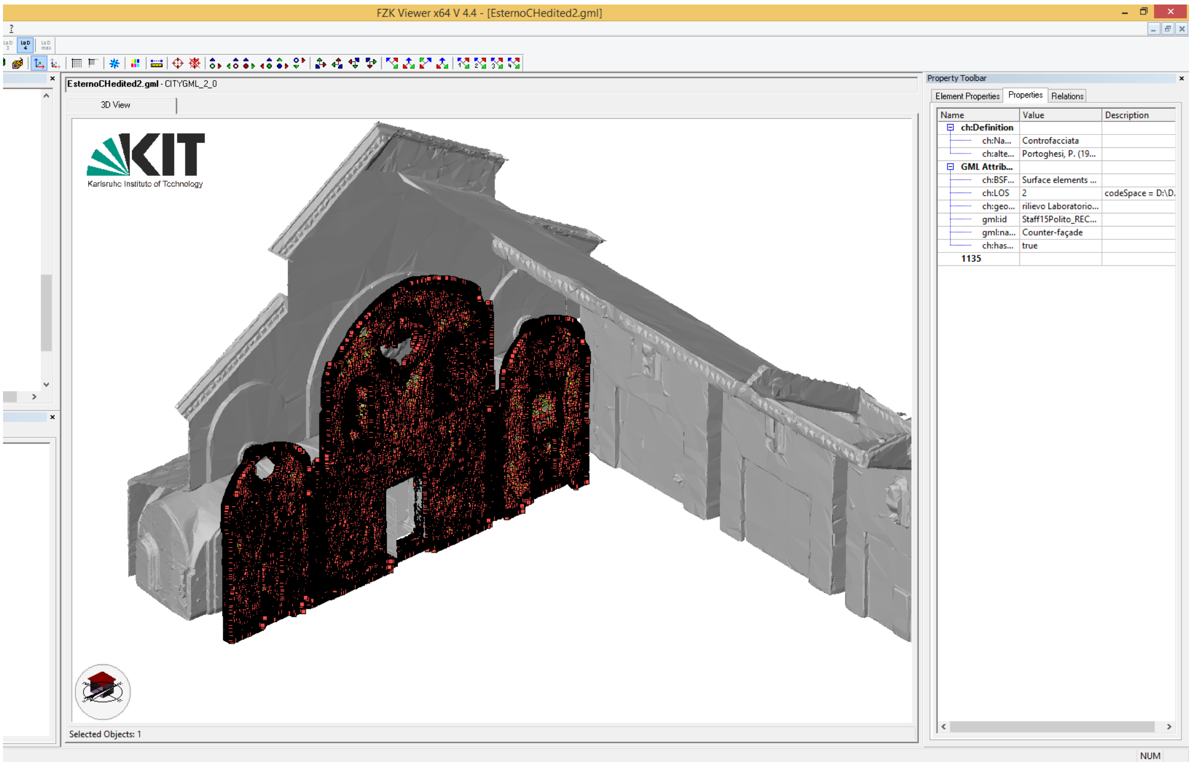


Figure 128. A sort of entity table can be opened; by selecting one of the entities, the geometry is selected and the properties are listed.

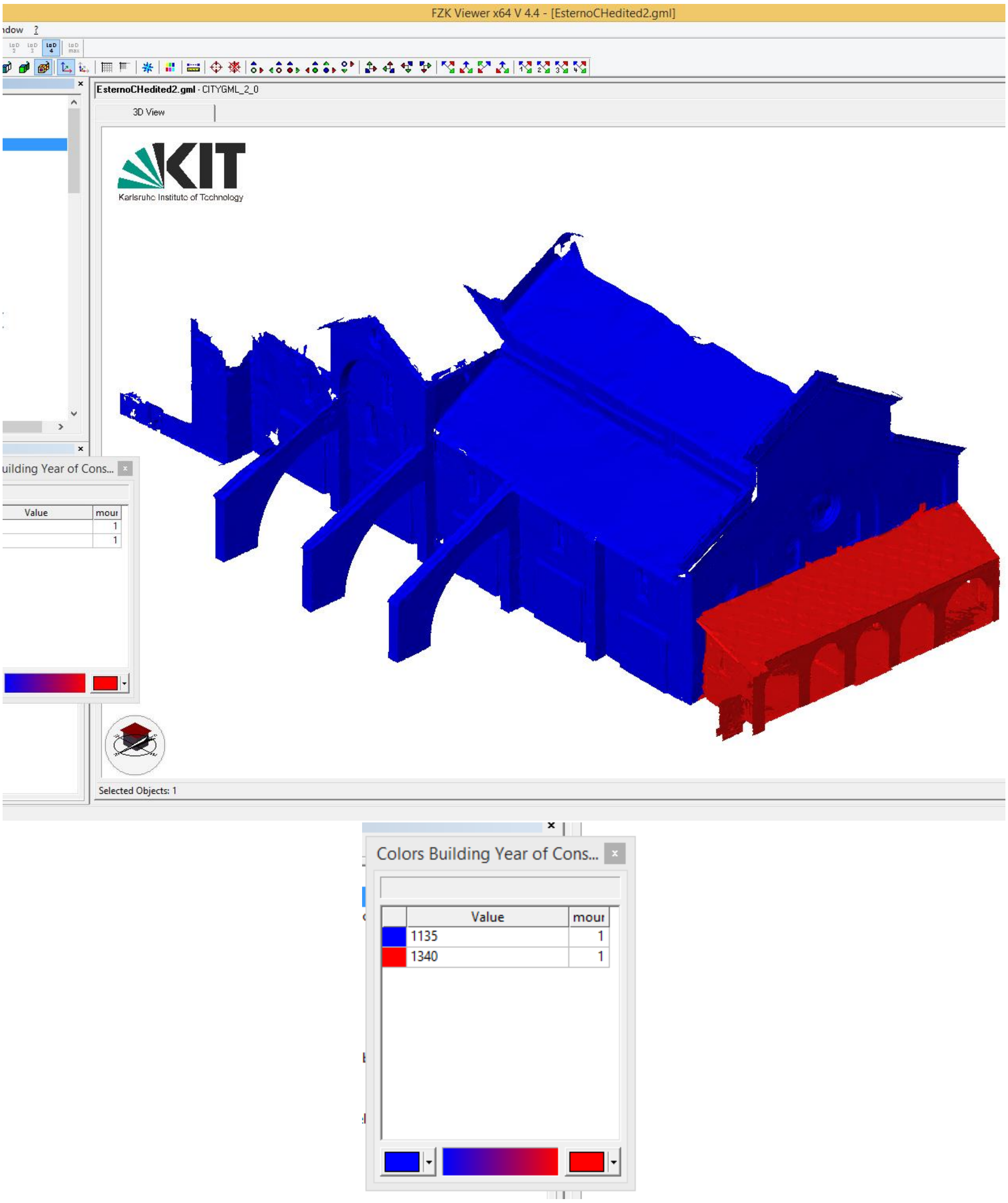


Figure 129. Example of thematic visualization based on the year of construction.

Some problems obviously remain, since both the XML schema and the software should be tested and optimized by computer scientists for enabling all the possibilities included in the implementation-independent model. Moreover, the tools and software applications were developed for representing urban-scale objects; therefore, some characteristics must be adapted and changed for better representing the architectonic level.

In summary, the potentialities of the system, especially if considered as part of a complete archive including all the resources in the Semantic Web, are clear.

11 AN ALTERNATIVE USE OF THE CITYGML CHADE MODEL: THE GENERATION OF AN SQL DATABASE

From the XSD file, SQL (Simple Query Language) relational or object-relational database schemas can also be generated, by passing through different software, such as *ALTOVA XMLSpy*, which permits the management of XML documents.

The process can generate the tables (a test has been done for the automatic generation by *ALTOVA XMLSpy* of a *PostgreSQL* database), but the simple translation is not effective, since a number of tables are generated (one for each element of the schema). These are so numerous because in object-oriented structures each value that would correspond to an attribute of a table in a relational database is an independent object, so that it is treated as a table itself. Further processing and intermediate setting files should solve this problem, if the community will consider this possibility to be a useful issue.

Some studies have been done about this storage of object-oriented ontologies into relational databases. An attempt was made by (Astrova et al., 2007) for the storage of OWL ontologies in SQL databases. A number of problems were encountered, among which the loss of structure would be the most serious.

Also, the Italian Spatial Data Group, working on the transformation of national, geographic databases to be INSPIRE-compliant, developed some tools to generate an SQL database from the INSPIRE UML structure (Belussi et al., 2014), but in this case something in the structure is also unavoidably lost.

The problem could potentially be solved by using a double structure for having the advantage of both systems in different parts of the archive (possibly, only as hypothesis, the static data should be object-oriented and the dynamic data should follow a relational model, such as monitoring data).

12 THE FURTHER STEP: THE TRANSLATION OF THE GML DATA TO LINKED OPEN DATA

The passage to linked open data would be the next step for achieving the maximum exploitation of the Semantic Web tools.

This requires two passages: the translation of GML data to RDF format and the translation of XSD to OWL.

The second one consists in changing the implementation level of the ontology, which remains the same. Some studies propose to pass through the UML model (that is, to start from the proper ontology) to generate a different schema in OWL format (van den Brink et al., 2014).

The first translation instead can be performed using XSLT and the use of some more complex structures, for example defined by geoSPARQL (van den Brink et al., 2014).

Another work for translating geographic data in typical geographic data formats to linked data is (Patroumpas et al., 2014), which presents a specific tool (*TripleGeo*) for performing this transformation.

The passage to linked open data, besides having the advantages provided by the technology itself, would permit the relation of the model and the data to further ontologies, such as the CIDOC CRM, for achieving a full representation of architectural heritage.

SECTION 4 – CONCLUSIONS AND PERSPECTIVES

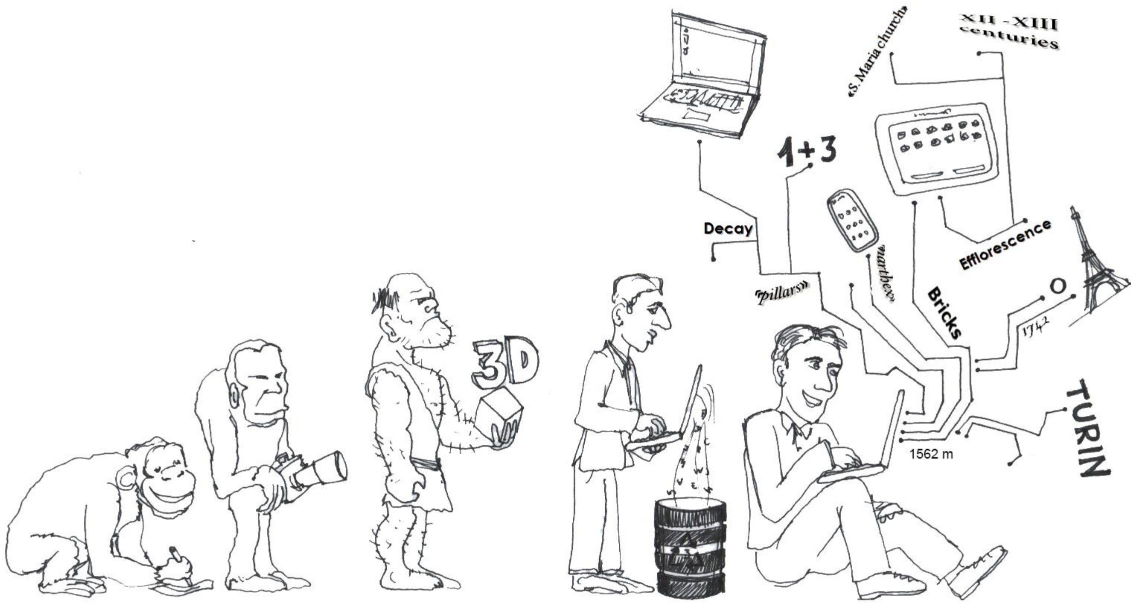


Image by E. Anastasi

13 A SYSTEM TOWARDS DOCUMENTATION INTELLIGENCE

Architectural heritage documentation has always been acknowledged to be essential for preservation and promotion. Moreover, the communication of the archived information is of a primary importance in order to foster identity issues, societal development, and the use of architectural heritage in a responsible and sustainable way.

It is important that the archived knowledge be correctly interpreted for being effectively understood, shared, and used.

For this aim, the development of standards for the documentation is critical. This need has been intensified since the development of digital technologies. This could be seen as a risk to the richness and multiplicity of architectural heritage documentation. However, the proper use of such standards for producing, storing, and sharing information can greatly enhance the power of the information and offer the possibility to enhance the quantity of information by performing inferences on datasets and improving their quality through the possibility to constrain and define the meaning the data carry. The information retrieval on the web has great advantage of that, and the interoperability of information for the reuse and exchange of datasetss become possible.

Standards, followed in more recent years by the Semantic Web, initially worked separately and, for a period of years, in an increasingly integrated framework for providing tools and concepts to make these requirements for documentation a reality.

The spatial information is a fundamental part of documentation for architectural heritage. It can carry a number of forms of information about (from lower to higher scales): the relationship of the monument with the landscape and with its urban or natural environment, from which function, role, and use can be understood; the distribution of the spaces; the spatial composition and modules; the building techniques; decorations; deterioration processes and so on.

Many techniques for producing spatial information exist and can deliver highly accurate, dense, and detailed 3D models. They can be used for the analysis and communication of architectural heritage. For an effective understanding and interpretation of architecture, it is necessary to consider the modularity of its composition, the relation among the parts, and hierarchies that structure them. In addition, their proportion and reciprocal position is meaningful for the interpretation by architecture experts (historians, building scientists, restorers, and so on). All these concepts are studied by researchers who traditionally deal with geographical topics, and they can be transposed to the field of higher-scale 3D models for responding to the cited requirements. Equally, the structures built in the geographical field for the sharing of data on the web and standards fostering interoperability can be translated to the same 3D models.

This research has dealt with the enrichment of such architectural heritage 3D models (high level of detail, high accuracy, high density) by means of the semantics defined by standards dealing with both spatial information and cultural heritage.

The proposed method starts from the existing model CityGML, in particular, the module ‘Building’, which is extended for including the possibility to store some elements of complexity essential for architectural heritage: the possibility to divide surfaces into smaller parts, and the inclusion of further attributes necessary to analyse cultural heritage.

This is developed following existing guidelines and procedures for building CityGML ADEs. The process is however not fully automated, notwithstanding the existence of algorithms and systems implemented in software. The lack of some pieces in the setting files or some inaccuracies in the schema writing may cause problems in the generated model. A manual editing of the XML is therefore needed for correcting the schema definition. Nevertheless, once executed the manual correction, the schema works well and can be used for structuring the data.

The final results of the data management through the composed CHADE are effective, since the model can be visualized, queried, and measured, and its attributes can be read, by means of an open source software.

However, the result is reached after a series of processing phases on a 3D model that is already supposed to be processed as a surface and fully optimized.

The reduction of the model edges is a first necessity, in order to reduce the computational weight without losing important details.

The second step is the progressive segmentation of the model, for which two aspects should be considered: the semantic definition and the correct interpretation of edges (of the 3D shape or given by the discontinuity of textures or patterns; for example, for decay mapping). The automation of the process is to be pursued. Some automatic systems exist, but they are mainly based on B-Rep, while here the necessity is to use extensive representations.

The other weakness in the process is the translation from a 3D model format to CityGML – CHADE. Often, several passages are still required for maintaining the georeferencing aspects and the model to be translated into a CityGML format. The inclusion of the CHADE must be realized manually.

This process consequently requires more automation. Moreover, the development of software to be able to read and to analyse the so-conformed models should be a challenge. The same is true for graphical interfaces and user-friendly implementation, as well as for linked data on the web, which are not extensively developed yet.

A future development for the data model is to include mereo-topological relations and further spatial rules and constraints in the model to enable inferences and automatic recognitions.

The models and data should actually be published as linked open data, and some test of inferences should be made for maximally exploiting and understanding the great potential they have.

In particular, for what concerns future perspectives for the proper system for managing architectural heritage documentation by means of semantics, ontologies and web technologies, a main issue could be the further optimization of the CityGML CHADE model developed in this thesis. The second step should be its integration with the recently developed ontology for standing buildings as cultural heritage (the extension CRMBE of the CIDOC CRM), which also exploits linked data technologies.

REFERENCES

The bibliographic references are grouped into sections organised following a schema similar to the chapters' organization, since this is more adherent to the development of the research and also follows some chronological sense in the emerging concepts. Some references could be cited for more than one topic; for the classification, the preeminent one was considered. Internally to each section, the order of the references is alphabetical and the web site references are reported in the final parts.

All the web sites links were verified on 28/01/2016.

THE ARCHITECTURAL HERITAGE DOCUMENTATION AS A RESSOURCE

Eppich, R., Amel, C. (2007). *Recording, Documentation and Information Management for the Conservation of Heritage Places: Illustrated Examples*. Los Angeles, CA: Getty Conservation Institute. http://hdl.handle.net/10020/gci_pubs/recordim_vol2

European Commission, Directorate General for Research and Innovation (2015). *Getting cultural heritage to work for Europe. Report of the Horizon 2020 Expert Group on Cultural Heritage.* Luxembourg: Publication Office of the European Union. KI-01-15-128-EN-N ISBN: 978-92-79-46046-3 DOI: 10.2777/745666. http://bookshop.europa.eu/it/getting-cultural-heritage-to-work-for-europe-pbKI0115128/;pgid=Iq1Ekni0.1lSR0OOK4MycO9B0000OzsW1T43;sid=1x3GNPnS5vrGNq8LiKRYk5v30M7A5_4B0c4=?CatalogCategoryID=ANIKABstUgUAAAEjCJEY4e5L

ICOMOS (1996). *Principles for the recording of monuments, groups of buildings and sites*, 11th ICOMOS General Assembly, Sofia, Bulgaria, from 5 to 9 October 1996.

ICOMOS (2007). *The ICOMOS Charter for the Interpretation and Preservation of Cultural Heritage Sites* http://www.enamecharter.org/downloads/ ICOMOS_Interpretation_Charter_EN_10-04-07.pdf

Jigyasu, R., Murthy, M., Boccardi, G., Marrion, C., Douglas, D., King, J., O'Brien, G., Dolcemascolo, G., Kim, Y., Albrito, P., Osihn, M. (2013). *Heritage and Resilience: Issues and Opportunities for Reducing Disaster Risks.* http://nrl.northumbria.ac.uk/17231/

Letellier, R., Schmid, W., LeBlanc, F., Eppich, R. (2007). *Recording, documentation, and information management for the conservation of heritage places*. J. Paul Getty Trust, Shaftesbury: Donhead. ISBN: 978-0-89236-925-6. http://philipmarshall.net/Teaching/rwuhp175/Letellier_2008_Guiding_Principles_Getty_Guidance.pdf

Panzeri, M. (2009). *Fonti, metafonti e GIS per l'indagine della struttura storica del territorio*. Celid.

Pickard, R. (2002). *European Cultural Heritage* (Vol. 1). Council of Europe.

Stylianidis, E., Patias, P., Quintero, M. (2011). *CIPA Heritage documentation: best practices and applications*, CIPA, Vol. XXXVIII-5/C19 ISSN:1682-1750.

UNESCO (2013). *Operational Guidelines for the Implementation of the World Heritage Convention*, revised 2013 http://whc.unesco.org/en/criteria/.

UNESCO (1972). *World Heritage Convention*. http://whc.unesco.org/en/convention/

UNESCO (2003). *Convention for the Safeguarding of Intangible Cultural Heritage*. http://www.unesco.org/culture/ich/index.php?lg=en&pg=00006.

United Nation General Assembly (2012). *The future we want.* Resolution adopted by the General Assembly on 27 July 2012, A/RES/66/288. https://documents-dds-ny.un.org/doc/UNDOC/GEN/N11/476/10/PDF/N1147610.pdf?OpenElement website: http://www.uncsd2012.org/thefuturewewant.html Accessed on 04/03/2016.

UNISDR, (2015). *Sendai Framework for disaster Risk Reduction 2015-2030.* http://www.unisdr.org/we/inform/publications/43291

ICOM

http://archives.icom.museum/objectid/heritage/fore.html

THE CHARACTERISTICS OF ARCHITECTURAL HERITAGE DOCUMENTATION

Almagro, A., Carbonara, G., Casiello, S., Coppo, D., Cundari, C., de Fiore, G., Docci, M., Fondelli, Kirova, T., Mandelli, E., Rota, P., Doria, R., Robotti, C., Saint-Aubin J.P., Starace, F. (1999). *Verso la "Carta del rilievo architettonico". Testo di base per* *la definizione dei temi.* Seminario Internazionale di Studio “Gli strumenti di conoscenza per il progetto di restauro”, Valmontone, settembre 1999.

Hemmleb, M., Weritz, F., Maierhofer, C. (2005). Damage detection on buildings surfaces with multi-spectral techniques. *CIPA 2005 XX International Symposium, Torino, Italy*.

FURTHER ELEMENTS FOR SPATIAL KNOWLEDGE

Allen, J. (1991). Time and time again: the many ways to represent time. *International Journal of Intelligent Systems* 6, pp. 341–355.

Berjawi, B., Chesneau, E., Duchateau, F., Favetta, F., Cunty, C., Miquel, M., Laurini, R. (2014). Representing uncertainty in visual integration. *Proceedings of the 20th International Conference on Distributed Multimedia Systems, Pittsburgh, USA,* pp. 365-372.

Berjawi, B., Duchateau, F., Chesneau, É., Favetta, F., Seccia, G., Cunty, C., Miquel M. Laurini, R. (2014). Uncertainty visualization of multi-providers cartographic integration. *Journal of Visual Languages & Computing*, *25*(6), pp. 995-1002.

Billen, R., Zlatanova, S. (2003). 3D spatial relationships model: a useful concept for 3D cadastre?. *Computers, Environment and Urban Systems*, *27*(4), pp. 411-425.

Brahim, L., Okba, K., Laurini, R. (2015). Mathematical framework for topological relationships between ribbons and regions. *Journal of Visual Languages & Computing*, *26*, pp. 66-81.

Bucher, B., Falquet, G., Clementini, E., Sester, M. (2012). Spatial relations and properties for semantically enhanced 3D city models. *Proc. of the Workshop on 3D Issues in Urban and Environmental Systems, Madrid, Spain*.

De Luca, L., Véron, P., Florenzano, M. (2005). Semantic-based modelling and representation of patrimony buildings. *SVE Worksop towards Semantic Virtual Environments,* pp. 1-11.

De Luca, L., Busarayat, C., Stefani, C., Véron, P., Florenzano, M. (2007). An integrated framework to describe, analyze, document and share digital representations of architectural buildings. *VAST 2007: futures technologies to empower the heritage professionals,* pp. 24-27.

De Luca, L., Véron, P., Florenzano, M. (2007). A generic formalism for the semantic modeling and representation of architectural elements. *The Visual Computer*, *23*(3), pp. 181-205.

Egenhofer, M. (1991). Topology and Reasoning: Reasoning about Binary Topological Relations. In: Gunther, O., Schek H.-J. *Proceedings of the Second Symposium on Large Spatial Databases, LNCS in Advances in Spatial Databases*, Springer-Verlag, pp.141–160.

Egenhofer, M., Herring, J. (1990). A Mathematical Framework for the Definition of Topological Relationships. *Proceedings of the 4th International Symposium on Spatial Data Handling,* pp.803–813.

Ellul, C., Haklay, M. (2006). Requirements for Topology in 3D GIS. *Transactions in GIS*, *10*(2), pp. 157-175.

Hilaga, M., Shinagawa, Y., Kohmura, T., Kunii, T. L. (2001). Topology matching for fully automatic similarity estimation of 3D shapes. *Proceedings of the 28th annual conference on Computer graphics and interactive techniques,* ACM. pp. 203-212.

Hmida, H. B., Cruz, C., Boochs, F., Nicolle, C. (2013). From 9-IM Topological Operators to Qualitative Spatial Relations using 3D Selective Nef Complexes and Logic Rules for bodies. *arXiv preprint arXiv:1301.4992*.

Laurini, R. (2015). Geographic Ontologies, Gazetteers and Multilingualism. *Future Internet*, *7*(1), pp. 1-23.

Laurini, R. (2015b). Primi passi per la modellazione delle regole geospaziali. *Atti del Convegno ASITA (Associazioni Scientifiche per le Informazioni Territoriali e Ambientali), Lecco, Italia, 29-30/9-1/10/2015,* ISBN 978-88-041232-2-7, pp. 501-508.

Laurini, R. (2014). A conceptual framework for geographic knowledge engineering. *Journal of Visual Languages & Computing*, *25*(1), pp. 2-19.

Laurini, R. (2012). Importance of spatial relationships for geographic ontologies. *Seventh International Conference on Informatics and Urban and Regional Planning INPUT,* pp. 122-134.

Laurini, R., Servigne, S. (2011). Potentialité du géoweb. *L'Espace géographique*, (2), pp. 109-116.

Lee, Y., Hsu, J. (1992). Spatial reasoning and similarity retrieval of images using 2D C-string knowledge representation, *Pattern Recognition*, 25 (3), pp. 305–318.

Leopold, J. L., Sabharwal, C. L., Ward, K. J. (2015). Spatial relations between 3D objects: The association between natural language, topology, and metrics. *Journal of Visual Languages & Computing*, *27*, pp. 29-37.

Mani, I., Hitzeman, J., Richer, J., Harris, D., Quimby, R., Wellner, B. (2008). SpatialML: Annotation Scheme, Corpora, and Tools. In *LREC*.

McMaster, B., Shea, S. (1992). *Generalization in Digital Cartography*. Association of American Geographers, Washington D.C.

Solovyov, S. A. (2012). Categorical foundations of variety-based topology and topological systems. *Fuzzy Sets and Systems*, *192*, pp. 176-200.

Stoter, J. E., Van Oosterom, P. J. M. (2002). Incorporating 3D geo-objects into a 2D geo-DBMS. *Proceedings FIG, ACSM/ASPRS, Washington DC, April 19-26, 2002*.

Wagner, D., Wewetzer, M., Bogdahn, J., Alam, N., Pries, M., Coors, V. (2013). Geometric-semantical consistency validation of CityGML models. *Progress and new trends in 3D geoinformation sciences* Springer Berlin Heidelberg, pp. 171-192.

Zlatanova, S., Rahman, A. A., Shi, W. (2004). Topological models and frameworks for 3D spatial objects. *Computers & Geosciences*, *30*(4), pp. 419-428.

Zlatanova, S. (2000). On 3D topological relationships. *Database and Expert Systems Applications, 2000. Proceedings. 11th International Workshop On,* IEEE, pp. 913-919.

THE LIFE OF ARCHITECTURAL HERITAGE INFORMATION

THE PRODUCTION

Bitelli; G., Gatta G. (2012). Georeferencing of an XVIII century technical map of Bologna (Italy). *E-PERIMETRON*, , 7/4, pp. 195 – 204

Brovelli, M., Crespi, M., Cilloccu, F., Dequal, S., Lingua, A. M. (2009). *Ortoimmagini 1: 10000 e modelli altimetrici-Linee guida.* CISIS.

Bryan, P., Blake, B. (2000). *Metric survey specifications for English Heritage*. English Heritage.

Guerra, F., Miniutti, D. (2000). The cartographic projections for the representation of double curved surfaces. *The International Archives of Photogrammetry and Remote Sensing, XIXth Congress of the International Society for Photogrammetry.*

Hurni, L., Sell, G. (2009). Cartography and architecture: interplay between reality and fiction. *The Cartographic Journal*, *46*(4), pp. 323-332.

Kauppinen, T., Mantegari, G., Paakkarinen, P., Kuittinen, H., Hyvönen, E., Bandini, S. (2010). Determining relevance of imprecise temporal intervals for cultural heritage information retrieval. *International journal of human-computer studies*, *68*(9), pp. 549-560.

Remondino, F., Rizzi, A. (2010). Reality-based 3D documentation of natural and cultural heritage sites—techniques, problems, and examples. *Applied Geomatics*, 2(3), pp. 85-100.

LIDAR TECHNIQUE: TERRESTRIAL LASER SCANNER FOR CH

Abdel, M. (2011). *3D laser scanners: history, applications, and future.*

Bonfanti, C., Chiabrando, F., Rinaudo, F. (2013). TLS data for architectural 2D representation and 3D modeling. Different approaches tested in the case of San Giovanni in Saluzzo (CN) Italy. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2, pp. 37-42.

Lerma, J. L., Navarro, S., Cabrelles, M., Villaverde, V. (2010). Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: the Upper Palaeolithic Cave of Parpalló as a case study. *Journal of Archaeological Science,* 37(3), pp. 499-507.

Fröhlich, C., Mettenleiter, M. (2004). Terrestrial laser scanning–new perspectives in 3D surveying. *International archives of photogrammetry, remote sensing and spatial information sciences,* 36(Part 8), W2.

Barton, J. (2009). 3D laser scanning and the conservation of earthen architecture: a case study at the UNESCO World Heritage Site Merv, Turkmenistan. *World Archaeology,* 41(3), pp. 489-504.

IMAGE-BASED METHODS: PHOTOGRAMMETRY AND CV ALGORITHMS

Gruen, A., Remondino, F., Zhang, L. (2004). 3D modeling and visualization of large cultural heritage sites at very high resolution: the Bamiyan valley and its standing Buddhas. *XXth ISPRS Congress "Geo-Imagery Bridging Continents,* pp. 12-23.

Snavely, N., Seitz, S. M., Szeliski, R. (2006). Photo tourism: exploring photo collections in 3D. In *ACM transactions on graphics (TOG)* 25(3), ACM, pp. 835-846.

Szeliski, R. (2010). *Computer vision: algorithms and applications*. Springer Science & Business Media.

UAV

Chiabrando, F., Lingua, A. M., Maschio, P. F., Rinaudo, F., Spanò, A. (2012). Mezzi aerei non convenzionali a volo autonomo per il rilievo fotogrammetrico in ambito archeologico. *Una giornata informale per i 70 anni del Prof. Carlo Monti* - 3 Maggio 2012, Milano, pp. 1-12.

Chiabrando F., Nex F., Piatti D., Rinaudo F. (2011). Uav And Rpv Systems For Photogrammetric Surveys. Archaelogical Areas: Two Tests In the Piedmont Region (ITALY). Journal of archaeological science, ISSN: 0305-4403, pp. 697-710.

Chiabrando, F., Marenchino, D., Nex, F., Piatti, D., Rinaudo, F. (2009). UAV systems for photogrammetric recording in archaeological areas: problems, suggestions, and products. *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology IX SPIE* SPIE Remote Sensing 2009, Berlin September 2009, ISBN: 9780819477835.

Colomina, I., Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *92*, pp. 79-97.

Nex, F., Remondino, F. (2014). UAV for 3D mapping applications: a review. *Applied Geomatics*, *6*(1), pp. 1-15.

Rau, J. Y., Chu, C. Y. (2012). Photo-Realistic 3D Mapping from Aerial Oblique Imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXVIII ISPRS Congress. Part* (Vol. 1).

Rinaudo, F., Chiabrando, F., Lingua, A. M., Spanò, A. T. (2012). Archaeological site monitoring: UAV photogrammetry can be an answer. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 39(B5), pp. 583-588.

Sauerbier, M., Eisenbeiss, H. (2010). UAVs for the documentation of archaeological excavations. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38(5), pp. 526-531.

Themistocleous, K., Ioannides, M., Agapiou, A., Hadjimitsis, D. G. (2015). The methodology of documenting cultural heritage sites using photogrammetry, UAV, and 3D printing techniques: the case study of Asinou Church in Cyprus. In Third International Conference on Remote Sensing and Geoinformation of the Environment. International Society for Optics and Photonics, pp. 953510-953510.

GEOREFERENCING METHODS

Aicardi, I., Chiabrando, F., Lingua, A., Noardo, F., Piras, M. (2016). UAV for thermal surveys as a support for watercourses and springs behavior description. *Geomatics, Natural Hazards and Risk.*

Barzaghi, R., Brovelli, M. A., Sona, G., Manzino, A., Sguerso, D. (1996). The new Italian quasigeoid: ITALGEO95. *Bollettino di geodesia e scienze affini*, *55*(1), pp. 57-72.

Brumana, R., Achille, C. (2007). E-historical sites. Georeferencing as availiability of space-temporal data: historical cartography towards advanced 3D view. *e-Perimetron*, 2(3), pp.134-159.

Brumana, R., Achille, C., Oreni, D., Prandi, F. (2005). From the guide of georeferencing cultural heritage for the map of risk (regione lombardia) to the georeferencing and troubleshooting in historical sites. *CIPA 2005 XX International Symposium*, 26 September – 01 October, 2005, Torino, Italy.

Brumana, R., Achille, C. Advanced GIS technologies to support georeferencing of the cultural heritage. *XIX CIPA Symposium. CIPA 2003 XIX International Symposium "New perspectives to save cultural heritage",* Antalya – Turkey, 30 September – 4 October 2003.

Chiabrando, F., Lingua, A., Piras, M. (2013). Direct photogrammetry using UAV: tests and first results. *ISPRS Archives*, *1*, W2.

Cina, A. (2014). *Dal GPS al GNSS (Global Navigation Satellite System)*, Celid, Torino. ISBN:978-88-6789-020-0

Cina, A., Baldo, M., Manzino A. (2000) GPS cinematico in appoggio a misure laser a scansione, *Atti della IV Conferenza ASITA*, Genova ottobre 2000, pp 113-122.

Fermi, M., Manzino, A., Vespe, F. (1998) la rete fiduciale di stazioni permanenti GPS in Italia, In: *bollettino della società italiana di topografia e fotogrammetria,* ISSN: 0392-4424, pp. 9-18.

Leick A, 2004. *GPS Satellite Surveying*. Wiley &Sons, New York.

Manzino AM, 2001, *Lezioni di topografia*, Otto editore, Torino.

Ventura, S. (2014). Linee guida per il geocoding e la georeferenziazione dei beni. *SIGECweb* 1.1 - marzo 2014. Referente ICCD: Antonella Negri, ICCD.

EPSG

www.epsg.org

METHODS AND MODELS INTEGRATION

Bastonero, P., Donadio, E., Chiabrando, F., Spanò, A. (2014). Fusion of 3D models derived from TLS and image-based techniques for CH enhanced documentation. *ISPRS Annals Of The Photogrammetry, Remote Sensing And Spatial Information Sciences*, *2*, pp. 73-80.

El-Hakim, S., Beraldin, J., Picard, M., Godin, G. (2004). Detailed 3D reconstruction of large-scale heritage sites with integrated techniques. *IEEE Computer Graphic Applications* 24(3), pp. 21–29.

GIS

Goodchild, M. F., Egenhofer, M. J., Fegeas, R. (1997). Interoperating GISs. *Report of a Specialist Meeting Held Under the Auspices of the Varenius Project Panel on Computational Implementations of Geographic Concepts.*

Laurini, R., Thompson, D. (1992). *Fundamentals of spatial information systems* (Vol. 37). Academic press.

Jafari Salim, M. (2014). *The Analysis of Open Source Software and Data for Establishment of GIS Services Throughout the Network in a Mapping Organization at National or International Level.* PhD thesis. Tutor: Piero Boccardo.

Worboys, M. F., Duckham, M. (2004). *GIS: a computing perspective*. CRC press.

GIS SPATIAL ANALYSIS

Fotheringham, S., Rogerson, P. (Eds.). (2013). *Spatial analysis and GIS*. CRC Press.

Maguire, D. J., Batty, M., Goodchild, M. F. (2005). *GIS, spatial analysis, and modeling*. ESRI Press.

Goodchild, M. F., Haining, R. P. (2004). GIS and spatial data analysis: Converging perspectives. *Papers in Regional Science*, *83*(1), pp. 363-385. DOI:10.1007/s10110-003-0190-y

DATABASE

Atkinson, M. P., Bancilhon, F., DeWitt, D. J., Dittrich, K. R., Maier, D., Zdonik, S. B. (1989). The Object-Oriented Database System Manifesto. In *DOOD,* Vol. 89. pp. 40-57.

Bloor, B. (2003). The failure of relational database, the rise of object technology and the need for the hybrid database. *Baroudi Bloor International Inc*, *2004*.

Codd, E. F. (1970). A relational model of data for large shared data banks. *Communications of the ACM*, *13*(6), pp. 377-387.

SQL

https://en.wikipedia.org/wiki/SQL

OMG

www.omg.org

UML

http://www.omg.org/gettingstarted/what_is_uml.htm
https://it.wikipedia.org/wiki/Unified_Modeling_Language
http://www.sparxsystems.com/enterprise_architect_user_guide/10/standard_uml_models/stereotypedlg.html

OBJECT ORIENTED APPROACH FOR GIS

Egenhofer, M. J., Frank, A. (1992). Object-oriented modeling for GIS. *Journal of the Urban and Regional Information Systems Association*, *4*(2), pp. 3-19.

Gaede, V., Riekert, W.F. (1994). Spatial access methods and query processing in the object-oriented GIS GODOT. *Proceedings of the AGDM'94 Workshop.*

Gong, J., Li, D. (1996). Design and implementation of an object-oriented GIS Software. *International archives of photogrammetry and remote sensing*, *31*, pp. 299-304.

Medeiros, C. B., Pires, F. (1994). Databases for GIS. *ACM Sigmod Record*, *23*(1), pp. 107-115.

Mennis J.L. (2003). Derivation and implementation of a semantic GIS data model informed by principles of cognition. *Computers, Environment and Urban Systems* 27, pp. 455-479.

Scholl, M., Voisard, A. (1992). Object-oriented database systems for geographic applications: an experiment with O2. *Proc. Int. Workshop on Database Management Systems for Geographical Applications,* Springer.

Worboys, M. F., Hearnshaw, H. M., Maguire, D. J. (1990). Object-oriented data modelling for spatial databases. *International journal of geographical information system*, *4*(4), pp. 369-383.

GIS CASE STUDIES FOR ARCHITECTURAL MANAGEMENT

Baracchini, C., Fabiani, F., Ponticelli, P., Vecchi, A. (2007). Verso un sistema unico di riferimento per la documentazione di restauro: storia e sviluppi di SICaR w/b. *Proceedings of eARCOM*, *2007*.

Box, P. (1999). *GIS and Cultural Resource Management: A manual for heritage managers.* UNESCO.

Centofanti, M., Continenza, R., Ruggieri, G., Brusaporci, S., Trizio, I. (2008). Il progetto del SIARCH-UNIVAQ, Sistema Informativo Architettonico. *DISEGNARECON*, *1*(2).

Cerutti, E., Noardo, F., Spanò, A. (2015). Architectural Heritage Semantic Data Managing and Sharing in GIS. *GISTAM 2015: 1st International Conference on Geographical Information System Theory, Application and Management, 27-30 April 2015, Barcelona (Spain)*. SciTePress.121-128 ISBN: 978-989-758-099-4 DOI: 10.5220/0005387801210128.

De Luca, L., Busayarat, C., Stefani, C., Véron, P., Florenzano, M. (2011). A semantic-based platform for the digital analysis of architectural heritage. *Computers & Graphics*, *35*(2), pp. 227-241.

Donadio E. and Spanò A. (2015). Data Collection and Management for Stratigraphic Analysis of Upstanding Structures. *Proceedings of the 1st International Conference on Geographical Information Systems Theory, Applications and Management*, ISBN 978-989-758-099-4, pp. 34-39. DOI: 10.5220/0005470200340039

Drap, P., Merad, D., Seinturier, J., Boï, J. M., Peloso, D., Vannini, G., Nucciotti, M., Pruno, E. (2012). An information system for medieval archaeology based on photogrammetry and archaeological database: the Shawbak castle project. *Progress in Cultural Heritage Preservation*. Springer Berlin Heidelberg, pp. 119-128

Drap, P., Seinturier, J., Chambelland, J. C., Gaillard, G., Pires, H., Vannini, G., Mucciotti, M., Pruno, E. (2009). Going to Shawbak (Jordan) and getting the data back: toward a 3D GIS dedicated to medieval archaeology. *Proceedings of 3D Arch*, pp. 28-31.

Guarnieri, A., Pirotti, F., Vettore, A. (2010). Cultural heritage interactive 3D models on the web: An approach using open source and free software. *Journal of Cultural Heritage*, 11(3), pp. 350-353.

Meyer, É., Grussenmeyer, P., Perrin, J. P., Durand, A., Drap, P. (2007). A web information system for the management and the dissemination of Cultural Heritage data. *Journal of Cultural Heritage*, *8*(4), pp. 396-411.

Katsianis, M., Tsipidis, S., Kotsakis, K., & Kousoulakou, A. (2008). A 3D digital workflow for archaeological intra-site research using GIS. *Journal of Archaeological Science*, *35*(3), pp. 655-667.

Petrescu, F. (2007). The use of GIS technology in cultural heritage. In *Proceedings of the XXI International CIPA Symposium, Athens, Greece,* pp. 01-06.

Salonia, P., Negri, A. (2003). Historical buildings and their decay: data recording, analysing and transferring in an ITC environment. *International archives of photogrammetry remote sensing and spatial information sciences*, *34*(5/W12), pp. 302-306.

Stefani, C., Brunetaud, X., Janvier-Badosa, S., Beck, K., De Luca, L., Al-Mukhtar, M. (2014). Developing a toolkit for mapping and displaying stone alteration on a web-based documentation platform. *Journal of Cultural Heritage*, *15*(1), pp. 1-9.

Wüst, T., Nebiker, S., & Landolt, R. (2004). Applying the 3D GIS DILAS to Archaeology and Cultural Heritage Projects Requirements and First Results. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, *35*(Part 5), pp. 407-412.

Arkis project

www.arkis.it
http://www.cnr.it/istituti/FocusByN.html?cds=098&nfocus=3

SICaR w/b

http://www.artpast.iccd.beniculturali.it/fileadmin/template/allegati/ rearte/SICaR_scheda_tecnica.pdf
http://sicar.beniculturali.it:8080/
http://www.sbappsae-pi.beniculturali.it/index.php?it/300/sistema-informativo-sicar
www.liberologico.com/sicar

BIM, HBIM AND CITYGML – BIM INTEGRATION

Azhar, S., Nadeem, A., Mok, J. Y., Leung, B. H. (2008). Building Information Modeling (BIM): A new paradigm for visual interactive modeling and simulation for construction projects. *Proc., First International Conference on Construction in Developing Countries.* pp. 435-446.

El-Mekawy, M., Östman, A., Hijazi, I. (2012). An evaluation of IFC-CityGML unidirectional conversion. *International Journal of Advanced Computer Science and Applications*, *3*(5). pp. 159-171.

El-Mekawy, M., Östman, A. (2010). Semantic mapping: an ontology engineering method for integrating building models in IFC and CityGML. *3rd ISDE digital earth summit.* pp. 12-14.

Hichri, N., Stefani, C., De Luca, L., Veron, P., Hamon, G. (2013). From point cloud to BIM: a survey of existing approaches. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *1*(2), pp. 343-348.

Level Of Development

http://bimforum.org/wp-content/uploads/2013/08/2013-LOD-Specification.pdf

3D MODELS AS A SUPPORT FOR DATA ANALYSIS

Chiabrando, F., Donadio, E., Sammartano, G., Spanò, A. (2015). La tecnologia laser scanning per la valutazione statica delle strutture storiche. *In: Conferenza Nazionale ASITA 2015, Lecco, 29 Settembre - 1 Ottobre 2015*. pp. 253-261.

Castellazzi, G.; D'Altri, A.M.; Bitelli, G.; Selvaggi, I.; Lambertini, A. (2015). From Laser Scanning to Finite Element Analysis of Complex Buildings by Using a Semi-Automatic Procedure. *Sensors*, *15*, pp. 18360-18380.

WEBGIS

Agosto, E., Dalmasso, S., Pasquali, P. (2011). Rapid WebGIS Development for Emergency Management. *Gi4DM (7th editioin), Antalya (Turkey), 3-7 May 2011.*

Agosto, E., Ardissone, P., Rinaudo, F. (2007). GIS And Web-GIS, Commercial and Open Source Platforms: General Rules for Cultural Heritage Documentation. *International archives of the photogrammetry, remote sensing and spatial information sciences*, vol. XXXVI-5/C53, pp. 625-630. - ISSN 1682-1750

Boroushaki, S., Malczewski, J. (2010). Participatory GIS: A Web-based Collaborative GIS and Multicriteria Decision Analysis. *Urisa Journal*, *22*(1), 23.

Dragićević, S., Balram, S. (2004). A Web GIS collaborative framework to structure and manage distributed planning processes. *Journal of Geographical Systems*, *6*(2), pp. 133-153.

Gallo, C., Malatacca, F. (2012). *Designing a WebGIS for spatial planning* (No. 01-2012). Dipartimento di Scienze Economiche, Matematiche e Statistiche, Universita'di Foggia.

Miler, M., Odobašić, D., Medak, D. (2010). Efficient Web-GIS Solution based on Open Source Technologies: Case-Study of Urban Planning and Management of the City of Zagreb, Croatia. *The XXIV FIG International Congress–Facing the Challenges–Building the Capacity*. Hrvatska znanstvena bibliografija i MZOS-Svibor.

Milosavljević, A., Stoimenov, L., Đorđević-Kajan, S. (2005). An architecture for open and scalable WebGIS. *8th AGILE Conference on GIScience, Estoril, Portugal.* pp. 629-634.

Pascaul, M., Alves, E., de Almeida, T., de França, G. S., Roig, H., Holanda, M. (2012). An Architecture for Geographic Information Systems on the Web–webGIS. *Proc. Fourth International Conference on Advanced Geographic Information Systems, Applications and Services, GEOProcessing*.

geoserver.org

geonode.org/
http://www.agire.it/

ONTOLOGIES

Baader, F., Nutt, W. (2003). Basic description logics. *Description logic handbook.* pp. 43-95.

Borgida, A., Serafini, L. (2003). Distributed description logics: Assimilating information from peer sources. *Journal on Data* Semantics I. Springer Berlin Heidelberg. pp. 153-184.

Borst, W. N. (1997). *Construction of engineering ontologies for knowledge sharing and reuse*. Universiteit Twente.

d'Aquin, M., Gangemi, A. (2011). Is there beauty in ontologies?. *Applied Ontology*, 6(3), pp. 165-175.

Fellbaum, C. (1998). WordNet. Blackwell Publishing Ltd.

Galton, AP. (2005). Ontology is not just about objects: Poorly individuated phenomena in geo-ontology, *Applied Ontology*, vol. 1, pp. 47-52.

Ghidini, C., Giunchiglia, F. (2001). Local models semantics, or contextual reasoning= locality+ compatibility. Artificial intelligence, 127(2), pp. 221-259.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. Knowledge acquisition, 5(2), pp. 199-220.

Guarino, N., Oberle, D., Staab, S. (2009). What is an ontology?. *Handbook on ontologies* (). Springer Berlin Heidelberg, pp. 1-17.

Guarino, N., Musen, M. A. (2005). Applied ontology: Focusing on content, *Applied Ontology* 1(1), pp. 1-5.

Guarino, N., Welty, C. (2002). Evaluating ontological decisions with OntoClean. *Communications of the ACM*, *45*(2), pp. 61-65.

Guarino, N., Welty, C. (2002b). Identity and subsumption. *The Semantics of Relationships*. Springer Netherlands. pp. 111-126.

Guarino, N., Welty, C. (2000). A formal ontology of properties. *Knowledge Engineering and Knowledge Management Methods, Models, and Tools*. Springer Berlin Heidelberg. pp. 97-112.

Guarino, N. (1994). The Ontological Level. In: Casati, R., Smith, B., White, G. (eds.), *Philosophy and the Cognitive Sciences*, Vienna: Hölder-Pichler-Tempsky.

Keet, C. M. (2006). Introduction to part-whole relations: mereology, conceptual modelling and mathematical aspects. *KRDB Research Centre Technical Report KRDB06-3 (Tutorial reader), Faculty of Computer Science, Free University of Bozen-Bolzano, Italy*, *2*.

Keita, A., Laurini, R., Roussey, C., Zimmerman, M. (2004). Towards an ontology for urban planning: The towntology project. *CD-ROM Proceedings of the 24th UDMS Symposium, Chioggia* (Vol. 12, No. 1).

Keller, U. (2004). *Some Remarks on the Definability of Transitive Closure in First-order Logic and Datalog*. Internal report, Digital Enterprise Research Institute (DERI), University of Innsbruck.

Mascardi, V., Briola, D., Locoro, A., Grignani, D., Deufemia, V., Paolino, L., Bianchi, N., de Lumley, H., Malafronte, D., Ricciarelli, A. (2014). A holonic multi-agent system for sketch, image and text interpretation in the rock art domain. *International Journal of Innovative Computing, Information and Control*. ISSN:1349-4198.

Masolo, C., Oltramari, A., Gangemi, A., Guarino, N., Vieu, L. (2003). La prospettiva dell'ontologia applicata. *Rivista di estetica*, *22*(2003), pp. 170-183.

Mulligan, K. (2013). Metafisica e ontologia. *Swiss Philosophical Preprint Series*, (44).

Mulligan, K. (1998). Relations–through thick and thin. *Erkenntnis*, *48*(2-3), pp. 325-353.

Nardi, D., Brachman, R. J. (2003). An Introduction to Description Logics. *Description logic handbook,* pp. 1-40.

Ogden, C. K., Richards, I. A. (1923). *The Meaning of Meaning: A Study of the Influence Of Language Upon Thought and of the Science of Symbolism*. Routledge & Kegan Paul Limited. London, tenth edition.

Smith, B., Ceusters, W. (2010). Ontological realism: A methodology for coordinated evolution of scientific ontologies. *Applied ontology*, *5*(3-4), 139.

Smith, B. (2003). Ontology. In: Floridi, L., *Blackwell guide to the philosophy of computing and Information*, Oxford: Blackwell, 2003, pp. 155-166.

Sowa, J. F. (2010). The role of logic and ontology in language and reasoning. *Theory and Applications of Ontology: Philosophical Perspectives*. Springer Netherlands. pp. 231-263.

Studer, R., Benjamins, V. R., Fensel, D. (1998). Knowledge engineering: principles and methods. *Data & knowledge engineering*, *25*(1), pp. 161-197.

Quine, W. V. (1953). On what there is. *From a logical point of view*, New York: Harper & Row.

Quine, W. V. (1948). On what there is. *The Review of Metaphysics*, *2*(1), pp. 21-38.

Spyns, P., Meersman, R., Jarrar, M. (2002). Data modelling versus ontology engineering. ACM SIGMod Record, 31(4), 12-17.

Vossen, P., Agirre, E., Calzolari, N., Fellbaum, C., Hsieh, S. K., Huang, C. R., Isahara, H., Kanzaki, K., Marchetti, A., Monachini, M., Raffaelli, R., Rigau, G., Tescon, M., VanGent, J., Neri, F. (2008). KYOTO: a System for Mining, Structuring and Distributing Knowledge across Languages and Cultures. In *LREC*.

DOLCE

https://en.wikipedia.org/wiki/Descriptive_Ontology_for_Linguistic_and_Cognitive_Engineering

BFO

http://ifomis.uni-saarland.de/bfo/

CYC

http://sw.opencyc.org/

KYOTO

http://weblab.iit.cnr.it/kyoto/xmlgroup.iit.cnr.it/kyoto/index8d0f.html?option=com_content&view=article&id=390&Itemid=155)

INDIANAMAS

indianamas.disi.unige.it

DOMAIN ONTOLOGIES IN THE MODELLING PROCESS OF AN INFORMATION SYSTEM

Benevides, A. B. (2010). *A model-based graphical editor for supporting the creation, verification and validation of OntoUML conceptual models* (Doctoral dissertation, Ph. D. thesis, Federal University of Espırito Santo (UFES), Vitória, ES, Brazil (Feb 2010)).

Guizzardi, G., Wagner, G. (2010). Using the unified foundational ontology (UFO) as a foundation for general conceptual modeling languages. In *Theory and Applications of Ontology: Computer Applications*, Springer Netherlands, pp. 175-196.

Guizzardi, G. (2005). *Ontological foundations for structural conceptual models*. PhD thesis. CTIT, Centre for Telematics and Information Technology.

Guizzardi, G., Wagner, G. (2004). A Unified Foundational Ontology and some Applications of it in Business Modeling. *CAiSE Workshops (3)* pp. 129-143.

Guizzardi, G. (2011). Ontological foundations for conceptual part-whole relations: the case of collectives and their parts. *Advanced Information Systems Engineering.* Springer Berlin Heidelberg, pp. 138-153.

Guizzardi, G. (2009). The problem of transitivity of part-whole relations in conceptual modeling revisited. *Advanced Information Systems Engineering*. Springer Berlin Heidelberg, pp. 94-109.

Guizzardi, G. (2010). On the Representation of Quantities and their Parts in Conceptual Modeling. *FOIS,* pp. 103-116.

Guizzardi, G., Wagner, G. (2008). What's in a relationship: an ontological analysis. *Conceptual Modeling-ER 2008*. Springer Berlin Heidelberg, pp. 83-97.

Costal, D., Gómez, C., Guizzardi, G. (2011). Formal semantics and ontological analysis for understanding subsetting, specialization and redefinition of associations in UML. *Conceptual Modeling–ER 2011.* Springer Berlin Heidelberg, pp. 189-203.

Benevides, A. B., Guizzardi, G., Braga, B. F. B., Almeida, J. P. A. (2010). Validating Modal Aspects of OntoUML Conceptual Models Using Automatically Generated Visual World Structures. *J. UCS*, *16*(20), pp. 2904-2933.

Guizzardi, G., das Graças, A. P., Guizzardi, R. S. (2011). Design patterns and inductive modeling rules to support the construction of ontologically well-founded conceptual models in OntoUML. *Advanced Information Systems Engineering Workshops*. Springer Berlin Heidelberg, pp. 402-413.

ONTOLOGY IN GIS

Chaturvedi, K., Yao, Z., Kolbe, T. H. (2015). Web-based Exploration of and Interaction with Large and Deeply Structured Semantic 3D City Models using HTML5 and WebGL*. DGPF Tagungsband 24/2015.*

Chaturvedi, K. (2014). *Web based 3D analysis and visualization using HTML5 and WebGL.* PhD thesis, University of Twente.

Du, H., Jiang, W., Anand, S., Morley, J., Hart, G., Leibovici, D., Jackson, M. (2011). An ontology based approach for geospatial data integration of authoritative and crowd sourced datasets. *Proceedings of the 25th International Cartographic Conference*.

Fallahi, G. R., Frank, A. U., Mesgari, M. S., Rajabifard, A. (2008). An ontological structure for semantic interoperability of GIS and environmental modeling. *International Journal of Applied Earth Observation and Geoinformation*, *10*(3), pp. 342-357.

Fonseca, F. T., Egenhofer, M. J., Davis, C. A., Borges, K. A. (2000). Ontologies and knowledge sharing in urban GIS. *Computers, Environment and Urban Systems*, *24*(3), pp. 251-272.

Fonseca, F., Egenhofer, M. (1999). Ontology-driven geographic information systems. *7th ACMsymposium on advances in geographic information systems. Kansas City, MO*, pp. 14-19.

Guarino, N. (1998). Formal ontology in information systems*: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy* (Vol. 46). Amsterdam, IOS press, pp. 3-15.

Iosifescu-Enescu, I., Hurni, L. (2007). Towards cartographic ontologies or “how computers learn cartography”. *23rd ICA conference,* pp. 4-10.

Jones, C. B., Abdelmoty, A. I., Fu, G. (2003). Maintaining ontologies for geographical information retrieval on the web. *On The Move to Meaningful Internet Systems 2003: CoopIS, DOA, and ODBASE*, Springer Berlin Heidelberg, pp. 934-951.

Laurini, R., Kazar, O. (2016). Geographic ontologies: survey and challenges. Journal for Theoretical Cartography, 9/2016. ISSN 1868-1387.

Mark D., Egenhofer, M., Hirtle, S., Smith, B. (2000). *UCGIS emerging research theme: Ontological foundations for geographic information science.*

Mennis, J. L. (2003). Derivation and implementation of a semantic GIS data model informed by principles of cognition. *Computers, Environment and Urban Systems*, *27*(5), pp. 455-479.

Schneider, T., Hashemi, A., Bennett, M., Brady, M., Casanave, C., Graves, H., Graves, H., Gruninger, M., Guarino, N., Levenchuk, A. Lucier, E., Obrst, L., Ray, S., Sriram, R.D., Vizedom, A., West, M., Whetzel, T., Yim, P. (2012). Ontology for big systems: The ontology summit 2012 communique. *Applied Ontology*, 7(3), pp. 357-371.

Zhao, T., Zhang, C., Wei, M., Peng, Z. R. (2008). Ontology-based geospatial data query and integration. In *Geographic Information Science*. Springer Berlin Heidelberg, pp. 370-392.

THE STANDARDISATION

CULTURAL HERITAGE STANDARDS

ICOMOS-ISCS, 2008. *Illustrated glossary on stone deterioration patterns, Monuments and sites,* ISBN : 978-2-918086-00-0

Mancinelli, M. L. (2015). *Standard per la catalogazione: principi generali,* ICCD.

Mancinelli, M. L. (2015a). MODI - *Modulo informativo,* ICCD.

Mancinelli, M. L., Negri, A. (2014). Verso gli open data: l’ICCD e gli standard nazionali. Alcune riflessioni per un quadro metodologico condiviso - giugno 2014*, ArcheoFOSS 2014 IX edizione 19-20 giugno 2014*, Verona.

ICOM

http://archives.icom.museum/objectid/heritage/intro2.html
http://archives.icom.museum/objectid/heritage/intro3.html
http://archives.icom.museum/objectid/heritage/getty.html
http://epoch-net.org/site/research/standards/

Heritage Documentation Programs

http://www.nps.gov/hdp/standards/index.htm
http://www.nps.gov/hdp/standards/HABS/HABSHistoryGuidelines.pdf
http://www.nps.gov/hdp/standards/HABS/ HABSDrawings.pdf
http://www.nps.gov/hdp/standards/PhotoGuidelines.pdf
http://www.nps.gov/hdp/
http://www.nps.gov/hdp/about.htm
http://www.nps.gov/hdp/standards/habsguidelines.htm

Core Data Index To Historic Buildings And Monuments Of The Architectural Heritage

http://archives.icom.museum/objectid/heritage/core.html
http://archives.icom.museum/objectid/heritage/intro3.html

Guarini Archives

http://www.regione.piemonte.it/cultura/guarini/index.htm

Italian documentation standards

http://www.cartadelrischio.it/
http://www.cartadelrischio.it/eng/info.html
http://www.catalogo.beniculturali.it/sigecSSU_FE/ Home.action?timestamp=1452186106974
www.culturaitalia.it
http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/documenti/progetto/sintesi_progettotecnicoscientifico.pdf
http://www.culturaitalia.it/opencms/documentazione_tecnica_it.jsp
http://www.iccd.beniculturali.it/index.php?it/376/il-sistema

European digital Cultural Heritage inventories

http://www.minervaeurope.org/
http://www.michael-culture.eu/
www.michael-culture.org
http://www.michael-culture.eu/about
http://www.michael-culture.eu/michael-multilingual-inventory
http://www.michael-culture.eu/athena-plus-646
http://www.europeana.eu/portal/
EPOCH

http://public-repository.epoch-net.org/reports/CI_Tool_interpretation_management_v7.pdf
http://public-repository.epoch-net.org/presentations/rome/Procedures%20for%20Data%20Integration%20through%20CIDOC%20CRM.pdf
http://epoch-net.org/site/research/newtons/ama/
http://epoch-net.org/site/mad/
http://epoch-net.org/site/research/standards/

DIGITAL GEOGRAPHIC INFORMATION STANDARDS

Costamagna, E., Spanò, A. (2013). CityGML for Architectural Heritage. *Developments in Multidimensional Spatial Data Models*. Springer Berlin Heidelberg, pp. 219-237.

Costamagna, E., Spanò, A. (2012). Semantic models for Architectural Heritage documentation. *Lecture Notes in Computer Science*, Springer, pp. 241-250.

De Laat, R., Van Berlo, L. (2011) Integration of BIM and GIS: The Development of the CityGML GeoBIM Extension. *Advances in 3D Geo-Information Sciences*. LNG&C, pp. 211-225.

Fan, H. C., Meng, L. Q. (2009). Automatic derivation of different levels of detail for 3D buildings modelled by CityGML. In *24th International Cartography Conference, Santiago, Chile,* pp. 15-21.

ISO (2015). *ISO/TC 211 Strategic business plan.* 2015/10/23 Version 5.

Lake, M. R., Burggraf, D., Trninic, M., Rae, L. (2004). *Geography mark-up language: foundation for the geo-web*. John Wiley & Sons.

OGC (2012). *OGC City Geography Markup Language (CityGML) Encoding Standard*. Approval Date: 2012-03-09, Publication Date: 2012-04-04, External identifier of this OGC® document: http://www.opengis.net/spec/citygml/2.0, Reference number of this OGC® project document: OGC 12-019, Version: 2.0.0. Editors: Gröger G., Kolbe T.H., Nagel C., Häfele K.H.

OGC (2012a). *CityGML UML diagrams as contained in CityGML Encoding Standard.* Version 2.0, OGC Doc. No. 12-019 Jointly developed by the Special Interest Group 3D (SIG 3D) and the OGC CityGML SWG, 2002 – 2012. www.citygml.org

OGC (2012b). *Gazetteer Service - Application Profile of the Web Feature Service Best Practice*, OGC Best Practice, 2012. version 1.0, doc. n. OGC 11-122r1

OGC (2007). *OpenGIS® Geography Markup Language (GML) Encoding Standard*. Open Geospatial Consortium Inc., 2007-08-27, Reference number of this OGC® document: OGC 07-036, Version: 3.2.1, Category: OpenGIS® Standard, Editor: Clemens Portele.

Kolbe, T. H., Gröger, G., Plümer, L. (2005). CityGML: Interoperable access to 3D city models. *Geo-information for disaster management*. Springer Berlin Heidelberg, pp. 883-899.

Van den Brink, L., Stoter, J. E., Zlatanova, S. (2012). Modelling an application domain extension of CityGML in UML. *ISPRS Conference 7th International Conference on 3D Geoinformation, The International Archivees on the Photogrammetry, Remote Sensing and Spatial Information Sciences,* Volume XXXVIII-4, part C26, 16–17 May 2012, Québec, Canada. ISPRS.

OGC STANDARDS

http://www.opengeospatial.org/standards/
http://www.opengeospatial.org/standards/gml
https://en.wikipedia.org/wiki/ Geography_Markup_Language
http://www.citygml.org/?id=1538.
http://www.opengeospatial.org/standards/citygml
http://www.citygmlwiki.org/index.php/CityGML-ADEs
http://www.citygmlwiki.org/index.php/ XML_namespaces_and_schemaLocations

IFC

https://en.wikipedia.org/wiki/Industry_Foundation_Classes

ISO/TC211

http://www.isotc211.org/

INSPIRE

http://inspire.jrc.ec.europa.eu/

SMART CITIES

Al-Hader, M., Rodzi, A., Sharif, A. R., Ahmad, N. (2009). Smart city components architecture. *Computational Intelligence, Modelling and Simulation, 2009. CSSim'09. International Conference on*. IEEE, pp. 93-97.

Amato, F., Chianese, A., Moscato, V., Picariello, A., Sperli, G. (2012). SNOPS: a smart environment for cultural heritage applications. *Proceedings of the twelfth international workshop on Web information and data management*. ACM, pp. 49-56.

Bowerman, B., Braverman, J., Taylor, J., Todosow, H., Von Wimmersperg, U. (2000). The vision of a smart city. *2nd International Life Extension Technology Workshop, Paris* (Vol. 28).

Nam, T., Pardo, T. A. (2011). Smart city as urban innovation: Focusing on management, policy, and context. *Proceedings of the 5th International Conference on Theory and Practice of Electronic Governance*. ACM, pp. 185-194.

Nam, T., Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*. ACM, pp. 282-291.

Prandi, F., De Amicis, R., Piffer, S., Soavea, M., Cadzowb, S., Boix, E. G., D'Hondt, E. (2013). Using CityGML to deploy smart-city services for urban ecosystems. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *4*, W1.

Ryan, N., Cinotti, T. S., Raffa, G. (2005). Smart environments and their applications to cultural heritage. *Smart Environments and their Applications to Cultural Heritage*, 7.

Su, K., Li, J., Fu, H. (2011). Smart city and the applications. *Electronics, Communications and Control (ICECC), 2011 International Conference on*. IEEE, pp. 1028-1031.

Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M. (2014). Internet of things for smart cities. *Internet of Things Journal, IEEE*, *1*(1), pp.22-32.

THE SEMANTIC WEB

Antoniou, G., Van Harmelen, F. (2004). *A semantic web primer*. MIT press.

Domingue, J. (2011). *The future Internet-achievements and technological promises. Future Internet assembly.*

Roman, D., Keller, U., Lausen, H., de Bruijn, J., Lara, R., Stollberg, M., Polleres, A., Feier, C., Bussler, C., Fensel, D. (2005). Web Service Modeling Ontology. *Applied ontology*, *1*(1), pp.77-106.

Waters, J., Powers, B. J., Ceruti, M. G. (2009). Global interoperability using semantics, standards, science and technology (GIS^{3}T). *Computer Standards & Interfaces*, *31*(6), pp.1158-1166. DOI:10.1016/j.csi.2009.03.001.

Zamborlini, V., Guizzardi, G. (2010). On the representation of temporally changing information in OWL. *Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2010 14th IEEE International*. IEEE, pp. 283-292.

W3C STANDARDS

www.w3.org
https://www.w3.org/standards/semanticweb/data
https://www.w3.org/2004/02/skos/
https://www.w3.org/Addressing/URL/uri-spec.html
https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
https://www.w3.org/TR/2008/REC-xml-20081126/
https://www.w3.org/XML/
http://www.w3schools.com/xml/
https://en.wikipedia.org/wiki/XML_Schema_(W3C)
https://www.w3.org/TR/xmlschema-1/
http://www.w3schools.com/xml/xml_validator.asp
http://www.w3schools.com/xsl/xpath_intro.asp
http://www.w3schools.com/xsl/xquery_intro.asp
http://www.w3schools.com/xml/ xml_xlink.asp
https://www.w3.org/TR/xslt
https://www.w3.org/International/O-URL-and-ident.html
https://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/
https://www.w3.org/TR/2012/REC-owl2-overview-20121211/
https://www.w3.org/TR/rdf-schema/
https://www.w3.org/TR/rdf-sparql-query/
https://www.w3.org/2001/sw/wiki/OWL

OPEN STAND

https://open-stand.org/about-us/principles/

LINKED OPEN DATA GETTY INSTITUTE

http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf

OPEN SOURCE COMMUNITY AND OPEN SOURCE SOFTWARE

Agenzia per l'Italia Digitale – Presidenza del Consiglio dei Ministri, Circolare 6 dicembre 2013 n.63 - Linee guida per la valutazione comparativa prevista dall'art. 68 del D.Lgs. 7 marzo 2005 n. 82 "*Codice dell'amministrazione digitale*".

Hars, A., Ou, S. (2001). Working for free? Motivations of participating in open source projects. *System Sciences, 2001. Proceedings of the 34th Annual Hawaii International Conference on*. IEEE.

Von Krogh, G., Spaeth, S., Lakhani, K. R. (2003). Community, joining, and specialization in open source software innovation: a case study. *Research Policy*, *32*(7), pp. 1217-1241.

Stallman, R. (1985). *The GNU manifesto*.

DOMAIN ONTOLOGIES IN THE SEMANTIC WEB

DiGiuseppe, N., Pouchard, L.C., Noy, N., (2014). SWEET ontology coverage for earth system sciences. *Earth Sci Inform*. DOI: 10.1007/s12145-013-0143-1.

Raskin, R. G., Pan, M. J. (2005). Knowledge representation in the semantic web for Earth and environmental terminology (SWEET). *Computers & geosciences*, *31*(9), pp. 1119-1125.

SWEET

https://sweet.jpl.nasa.gov/

GEONAMES

http://www.geonames.org/
http://www.geonames.org/ontology/documentation.html
https://en.wikipedia.org/wiki/GeoNames

CULTURAL HERITAGE CLASSIFICATIONS AND INVENTORIES EXPLOITING THE SEMANTIC WEB TECHNOLOGIES

THE CIDOC CRM AND ITS EXTENSIONS

Belussi, A., Migliorini, S., Cipriano, P., Negri, M., Pelagatti, G. (2014). Design of the Data Transformation Architecture for the INSPIRE Data Model Browser. *Connecting a Digital Europe through Location and Place. Proceedings of the AGILE'2014 International Conference on Geographic Information Science, Castellón, June, 3-6, 2014.* ISBN: 978-90-816960-4-3

Blaško, M., Cacciotti, R., Křemen, P., Kouba, Z. (2012). Monument damage ontology. *Progress in Cultural Heritage Preservation*. Springer Berlin Heidelberg. pp. 221-230

Cacciotti, R., Valach, J., Kuneš, P., Čerňanský, M., Blaško, M., řemen, P. (2013). Monument damage Information System (MONDIS), An Ontological Approach to Cultural Heritage Documentation. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*(5), pp.55-60.

De Roo, B., Bourgeois, J., De Maeyer, P. (2013). Conservation of Past Times: Data Models for Ensuring the Future of our Heritage, *Online proceedings of the conference Built Heritage 2013 Monitoring Conservation and Management,* pp. 982-988.

Doerr, M., Hiebel, G., Eide, Ø. (2013). CRMgeo: Linking the cidoc crm to geosparql through a spatiotemporal refinement. *Institute of Computer Science, Tech. Rep. GR70013*.

Doerr, M., Bekiari, C., Kritsotaki, A., Hiebel, G., Theodoridou, M. (2014). Modelling Scientific Activities: Proposal for a global schema for integrating metadata about scientific observation. *Access and understanding–networking in the digital era: The 6th annual conference of CIDOC, the International Committee for Documentation of ICOM, Dresden, Germany*. http://www.cidoc2014.de/images/sampledata/cidoc/papers/E-2_Bekiari_paper.pdf

Doerr, M. (2009). Ontologies for cultural heritage. In *Handbook on Ontologies*. Springer Berlin Heidelberg. pp. 463-486

Doerr M. (2002). The CIDOC CRM: an ontological approach to semantic interoperability of metadata. *AI Magazine, Special Issue*, *24*(3), pp. 75-92.

Eide, Ø., Felicetti, A., Ore, C. E., D'Andrea, A., Holmen, J. (2008). Encoding cultural heritage information for the semantic web. procedures for data integration through CIDOC-CRM mapping. *Open Digital Cultural Heritage Systems Conference.*

Hernández, F., Rodrigo, L., Contreras, J., Carbone, F. (2008). Building a cultural heritage ontology for Cantabria. *Annual Conference of CIDOC,* pp. 1-14.

ISO (2006). *ISO 21127:2006, Information and documentation -- A reference ontology for the interchange of cultural heritage information*. http://www.iso.org/iso/catalogue_detail?csnumber=34424 revised in *ISO 21127:2014* http://www.iso.org/iso/home/store/catalogue_ics/catalogue_detail_ics.htm?csnumber=57832.

Kondylakis, H., Doerr, M., Plexousakis, D. (2006). Mapping language for information integration. *FORTH-ICS, Technical Report*.

LeBoeuf, P., Doerr, M., Ore, Ch-E., Stead, S. (2013). *Definition of the CIDOC Conceptual Reference Model*, ICOM/CIDOC, 5.1.2.

Ronzino, P., Niccolucci, F., Felicetti, A., Doerr, M. (2015). CRMBA a CRM extension for the documentation of standing buildings. *International Journal on Digital Libraries*, pp. 1-8.

Stead, S., Doerr, M. et al. (2015). *CRMinf: the Argumentation Model. An Extension of CIDOC-CRM to support argumentation.* Version 0.7 http://www.ics.forth.gr/isl/CRMext/CRMinf/docs/CRMinf0.5.pdf.

CIDOC CRM

http://www.cidoc-crm.org/official_release_cidoc.html
http://www.cidoc-crm.org/docs/How_to%20implement%20CRM_Time_in%20RDF.pdf

THE GETTY INSTITUTE

Harpring, P. (2010). Development of the Getty vocabularies: AAT, TGN, ULAN, and CONA. *Art Documentation*, pp. 67-72.

Harpring, P. (2010a). *Introduction to controlled vocabularies: terminology for art, architecture, and other cultural works*. Getty Publications.

Getty Institute

www.getty.edu

http://www.getty.edu/about/whoweare/mission.html

Getty Vocabularies

http://www.getty.edu/research/tools/vocabularies/Linked_Data_Getty_Vocabularies.pdf
http://www.getty.edu/research/tools/ vocabularies/Linked_Data_Getty_Vocabularies.pdf
http://www.getty.edu/research/tools/vocabularies/lod/index.html
http://vocab.getty.edu
http://www.getty.edu/research/tools/vocabularies/aat/index.html
http://www.getty.edu/research/tools/vocabularies/tgn/index.html
http://www.getty.edu/research/tools/vocabularies/ulan/index.html
http://www.getty.edu/research/tools/vocabularies/cona/index.html

ARCHES project

http://www.getty.edu/conservation/our_projects/field_projects/arches/arches_overview.html

OGC ANSWERS TO THE SEMANTIC WEB PROPOSALS

OGC (2012c). *OGC GeoSPARQL – A Geographic Query Language for RDF Data*. OGC Implementation Standard, Approval Date: 2012-04-27, Publication Date: 2012-09-10, External identifier of this OGC® document:

http://www.opengis.net/doc/IS/geosparql/1.0, Reference number of this OGC® project document: OGC 11-052r4, Version: 1.0. Editors: Perry, M., Herring, J.

GeoSPARQL

http://www.opengeospatial.org/standards/geosparql

Geospatial ontologies

http://ows10.usersmarts.com/ows10/ontologies/

GeoKnow

http://geoknow.eu/Welcome.html

SEMANTIC SEGMENTATION AND EDITING OF 3D MODELS

Apollonio, I., Gaiani, M., Manfredini, A. (2010). Modellazione semantica metodi a multirisoluzione. In: Benedetti, B., Gaiani, M., Remondino, F., *Modelli digitali 3D in archeologia: il caso di Pompei*, ed. della Normale ISBN 978-88-7642-353-6, pp. 236-269.

Attene, M., Robbiano, F., Spagnuolo, M., Falcidieno, B. (2009). Characterization of 3D shape parts for semantic annotation. *Computer-Aided Design*, *41*(10), pp. 756-763.

Dorninger, P., Nothegger, C. (2007). 3D segmentation of unstructured point clouds for building modelling. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol 35, pp 191–196.

ESRI (2008). *The multipatch geometry type*. An ESRI white paper, ESRI

Golovinskiy, A., Funkhouser, T. (2009). Min-Cut Based Segmentation of Point Clouds, Computer Vision Workshops (ICCV Workshops), IEEE 12th International Conference, pp 39-46.

Ning, X., Zhang, X., Wang, Y, Jaeger, M., (2010). Segmentation of architecture shape information from 3D point cloud. In: Pottmann, H., Wallner, J., 1999. Approximation algorithms for developable surfaces. *Computer Aided Geometric Design*, Vol 16, pp. 539-556.

Zhana, Q., Liangb, Y., Xiaoa, Y. (2009). Color-based segmentation of point clouds . *IAPRS Vol XXXVIII*, Part 3/W8.

COLLADA

https://www.khronos.org/collada/

THE FURTHER STEP: THE TRANSLATION OF THE GML DATA TO LOD

Patroumpas, K., Alexakis, M., Giannopoulos, G., Athanasiou, S. (2014). TripleGeo: an ETL Tool for Transforming Geospatial Data into RDF Triples. *EDBT/ICDT Workshops,* pp. 275-278.

Van den Brink, L., Janssen, P., Quak, W., Stoter, J. E. (2014). Linking spatial data: automated conversion of geo-information models and GML data to RDF. *International Journal of Spatial Data Infrastructures Research 9,(2014)*.

CASE STUDIES USED FOR EXAMPLES

Staffarda Abbey (CN, Italy)

Bastonero, P. (2013). *UAV (Unmanned Aerial Vehicle): esperienza sull'Abbazia di Staffarda*. Rel. Spanò, A.T., Chiabrando, F. Politecnico di Torino, II Facoltà di architettura, Corso di laurea in Architettura per il progetto sostenibile, Tesi di Laurea Specialistica. Official URL: http://opac.biblio.polito.it:80/F/?func=direct&loc...

Beltramo, S. (2010). *L'abbazia cistercense di Santa Maria di Staffarda.* Architettura dei monasteri in Piemonte, L'artistica editrice. ISBN: 978-88-7320-262-2.

Donadio, Elisabetta (2013). *Tecnica laser scanning e tecnologia GIS per l'analisi stratigrafica del patrimonio storico costruito : la chiesa di Santa Maria dell'Abbazia di Staffarda.* Rel. Spanò, A.T., Beltramo, S., Chiabrando, F. Politecnico di Torino, II Facoltà di architettura, Corso di laurea in Architettura per il progetto sostenibile, Tesi di Laurea Specialistica. Official URL: http://opac.biblio.polito.it:80/F/?func=direct&loc...

Piedmont gypsum ceilings (Italy)

Noardo, F., Spanò, A. (2015). Towards a spatial semantic management for the Intangible Cultural Heritage. *International Journal of Heritage in the Digital Era*, Edition 2015/ Vol. 4 n.2, pp-133-147. DOI: 10.1260/2047-4970.4.2.133

Spanò, A.T. (2013). *Solai di gesso piemontesi*. CIRAAS, pp. 607, ISSN: 2037-9285 (collana CIRAAS), ISBN: 9788890624315

Stupinigi royal residence (TO, Italy)

Spanò A., Chiabrando F. (2013). *Report del Rilievo laser scanning ed elaborazioni orto fotogrammetriche del Salone Centrale della Palazzina di Caccia di Stupinigi*. Politecnico di Torino, DAD.

Chiabrando, F., Lingua, A., Noardo, F., Spanò, A. (2014). 3D Modelling of trompe l'oeil decorated vaults using dense matching techniques. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*(5), 97.

Turin roofs (TO, Italy)

Chiabrando, F., Danna, C., Lingua, A., Noardo, F., Osello, A. (2016). Effective tools for supporting energetic policies: 3d urban models generation and analysis. *International Journal Of Spatial Data Infrastructures Research*, 11(2016). ISSN 1725-0463.

Susa monuments (TO, Italy)

Chiabrando, F., Donadio, E., Fernandez-Palacios, B., Remondino, F., Spanò, A. (2015). L'arco di Augusto a Susa: un nuovo modello digitale per rinnovate indagini. In: *L'arco di Susa e i monumenti della propaganda imperiale in età augustea. Atti del Convegno di studi (Susa, 12 aprile 2014). Segusium - Soc. di Ricerche e Studi Valsusini,* pp. 217-232. ISBN 9788890785948

Aicardi, I., Boccardo, B., Chiabrando, F., Donadio, E., Lingua, A., Maschio, P., Noardo, F., Spanò, A. (2015). Modelli 3D multi sensore per l'acropoli segusina. In: *L'arco di Susa e i monumenti della propaganda imperiale in età augustea. Atti del Convegno di studi (Susa, 12 aprile 2014). Segusium - Soc. di Ricerche e Studi Valsusini*, pp. 217-232. ISBN 9788890785948

Colletto convent (TO, Italy)

Noardo, F. (2012). *Il progetto di conservazione: nuove tecnologie (laser scanner e GIS) per la sua formulazione: l'ex convento dei Carmelitani del Colletto*. Rel. Vinardi, M.G., Spanò, A.T., Politecnico di Torino, II Facoltà di architettura, Corso di laurea in architettura (restauro e valorizzazione), Tesi di Laurea Specialistica. Official URL: http://opac.biblio.polito.it:80/F/?func=direct&loc...

I’d like to thank my tutors, Andrea and Nannina, for proposing me such a challenging topic and for all what they do beyond their academic duties.

Thank you to Filiberto, Marco, Paolo, Paolo, Irene, Nives, Vincenzo (and Andrea and Nannina) for their continuous presence and help, and for making these hard-work years so enjoyable.

Thank you to Emilio Anastasi, who helped me with some images in the thesis.

Thank you to mamma, papà, Andrea, Fede (in a special way), Anna, Saverio, for always being there in any event.

Thank you to Michele and Rossella, for being my greatest achievement.

ANNEXES

Annex A A selection of OGC standards list

<table>
<tr><th>Document Title</th><th>Version</th><th>Doc.#</th><th>Editor</th><th>Date</th></tr>
<tr><td>Geographic information — Well known text representation of coordinate reference systems</td><td>1.0</td><td>12-063r5</td><td>Roger Lott</td><td>2015-05-01</td></tr>
<tr><td colspan="5">This Standard provides an updated version of WKT representation of coordinate reference systems that follows the provisions of ISO 19111:2007 and ISO 19111-2:2009. It extends the earlier WKT to allow for the description of coordinate operations. This International Standard defines the structure and content of well-known text strings. It does not prescribe how implementations should read or write these strings. The jointly developed draft has also been submitted by ISO TC211 for publication as an International Standard document. The version incorporates comments made during both the OGC Public Comment Period as well as the ISO ballot for DIS (ISO TC211 document N3750).</td></tr>
<tr><td>Geospatial eXtensible Access Control Markup Language (GeoXACML) Version 1 Corrigendum</td><td>1.0.1</td><td>11-017</td><td>Andreas Matheus, Jan Herrmann</td><td>2011-05-12</td></tr>
<tr><td colspan="5">The OpenGIS® Geospatial eXtensible Access Control Markup Language Encoding Standard (GeoXACML) defines a geospatial extension to the OASIS standard “eXtensible Access Control Markup Language (XACML)” [www.oasis-open.org/committees/xacml/]. This extension incorporates spatial data types and spatial authorization decision functions based on the OGC Simple Features[http://www.opengeospatial.org/standards/sfa] and GML[http://www.opengeospatial.org/standards/gml] standards. GeoXACML is a policy language that supports the declaration and enforcement of access rights across jurisdictions and can be used to implement interoperable access control systems for geospatial applications such as Spatial Data Infrastructures. GeoXACML is not designed to be a rights expression language and is therefore not an extension of the OGC GeoDRM Reference Model (Topic 18 in the OpenGIS® Abstract Specification [http://www.opengeospatial.org/standards/as]).</td></tr>
<tr><td>GeoXACML Implementation Specification - Extension A (GML2) Encoding GeoXACML extA</td><td>1.0</td><td>07-098r1</td><td>Andreas Matheus</td><td>2008-02-23</td></tr>
<tr><td colspan="5">This document defines an extension to the GeoXACML Implementation Specification, Verison 1.0 for the GML2 geometry encoding as specified in the GML2 standard.</td></tr>
<tr><td>GeoXACML Implementation Specification - Extension B (GML3) Encoding GeoXACML extB</td><td>1.0</td><td>07-099r1</td><td>Andreas Matheus</td><td>2008-02-23</td></tr>
<tr><td colspan="5">This specification is a normative extension to the GeoXACML core Implementation Specification. It defines the GML3 encoding for geometries.</td></tr>
<tr><td>Observations and Measurements - XML Implementation</td><td>2.0</td><td>10-025r1</td><td>Simon Cox</td><td>2011-03-22</td></tr>
<tr><td colspan="5">This standard specifies an XML implementation for the OGC and ISO Observations and Measurements (O&M) conceptual model (OGC Observations and Measurements v2.0 also published as ISO/DIS 19156), including a schema for Sampling Features. This encoding is an essential dependency for the OGC Sensor Observation Service (SOS) Interface Standard. More specifically, this standard defines XML schemas for observations, and for features involved in sampling when making observations. These provide document models for the exchange of information describing observation acts and their results, both within and between different scientific and technical communities.</td></tr>
</table>

OGC Augmented Reality Markup Language 2.0 (ARML 2.0)	21.0	12-132r4	Martin Lechner	2015-02-24
This OGC® Standard defines the Augmented Reality Markup Language 2.0 (ARML 2.0). ARML 2.0 allows users to describe virtual objects in an Augmented Reality (AR) scene with their appearances and their anchors (a broader concept of a location) related to the real world. Additionally, ARML 2.0 defines ECMAScript bindings to dynamically modify the AR scene based on user behavior and user input.				
OGC City Geography Markup Language (CityGML) Encoding Standard	2.0	12-019	Gerhard Gröger, Thomas H. Kolbe, Claus Nagel, Karl-Heinz Häfele	2012-04-04
CityGML is an open data model and XML-based format for the storage and exchange of virtual 3D city models. It is an application schema for the Geography Markup Language version 3.1.1 (GML3), the extendible international standard for spatial data exchange issued by the Open Geospatial Consortium (OGC) and the ISO TC211. The aim of the development of CityGML is to reach a common definition of the basic entities, attributes, and relations of a 3D city model. This is especially important with respect to the cost-effective sustainable maintenance of 3D city models, allowing the reuse of the same data in different application fields.				
OGC Filter Encoding 2.0 Encoding Standard - With Corrigendum	2.0.2	09-026r2	Panagiotis (Peter) A. Vretanos	2014-08-18
A fundamental operation performed on a set of data or resources is that of querying in order to obtain a subset of the data which contains certain desired information that satisfies some query criteria and which is also, perhaps, sorted in some specified manner. This International Standard defines an abstract component, named AbstractQueryExpression, from which other specifications can subclass concrete query elements to implement query operations. This International Standard also defines an additional abstract query component, named AbstractAdhocQueryExpresison, which is derived from AbstractQueryExpression and from which other specifications can subclass concrete query elements which follow a query pattern composed of a list of resource types to query, a projection clause specifying the properties of those resources to present in the result, a projection clause composed of predicates that define the subset of resources or data in the result set and a sorting clause indicating to order in which the results should be presented. This pattern is referred to as an ad hoc query pattern since the server is not aware of the query until it is submitted for processing. This is in contrast to a stored query expression, which is stored and can be invoked by name or identifier. This International Standard describes an XML and KVP encoding of a system-neutral syntax for expressing the projection, selection and sorting clauses of a query expression. The intent is that this neutral representation can be easly validated, parsed and then translated into some target query language such as SPARQL or SQL for processing.				
OGC GeoSPARQL - A Geographic Query Language for RDF Data	1.0	11-052r4	Matthew Perry and John Herring	2012-06-12
This standard defines a set of SPARQL extension functions [W3C SPARQL], a set of RIF rules [W3C RIF Core], and a core RDF/OWL vocabulary for geographic information based on the General Feature Model, Simple Features [ISO 19125-1], Feature Geometry and SQL MM.				
OGC® Geography Markup Language (GML) - Extended schemas and encoding rules	3.3	10-129r1	Clemens Portele	2012-02-07
The Geography Markup Language (GML) is an XML encoding in compliance with ISO 19118 for the transport and storage of geographic information modelled in accordance with the conceptual modelling framework used in the ISO 19100 series of International Standards and including both the spatial and non-spatial properties of geographic features.				

OGC® GML Application Schema - Coverages - GeoTIFF Coverage Encoding Profile wcs_geotiff	1.0	12-100r1	Stephan Meissl	2014-05-28
This Interface Standard is a profile of the OGC® GML Application Schema –Coverages version 1.0 [OC 09-146r2]. This document specifies the usage of the GeoTIFF data format for the encoding of GML coverages. This encoding is used by several OGC services like the Web Coverage Service (WCS) 2.0 Interface Standard – Core [OGC 09-110r4].				
OGC® GML in JPEG 2000 (GMLJP2) Encoding Standard Part 1: Core	2.0	08-054r4	Lucio Colaiacomo, Joan Masó, Emmanuel Devys	2014-09-23
This standard applies to the encoding and decoding of JPEG 2000 images that contain GML for use with geographic imagery. This document specifies the use of the Geography Markup Language (GML) within the XML boxes of the JPEG 2000 data format and provides an application schema for JPEG 2000 that can be extended to include geometrical feature descriptions and annotations. The document also specifies the encoding and packaging rules for GML use in JPEG 2000.				
OGC® IndoorGML	1.0	14-005r3	Jiyeong Lee, Ki-Joune Li, Sisi Zlatanova, Thomas H. Kolbe, Claus Nagel, Thomas Becker	2014-12-02
This OGC® IndoorGML standard specifies an open data model and XML schema for indoor spatial information. IndoorGML is an application schema of OGC® GML 3.2.1. While there are several 3D building modelling standards such as CityGML, KML, and IFC, which deal with interior space of buildings from geometric, cartographic, and semantic viewpoints, IndoorGML intentionally focuses on modelling indoor spaces for navigation purposes.				
OpenGIS Geography Markup Language (GML) Encoding Standard	3.2.1	07-036	Clemens Portele	2007-10-05
The OpenGIS® Geography Markup Language Encoding Standard (GML) The Geography Markup Language (GML) is an XML grammar for expressing geographical features. GML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on the Internet. As with most XML based grammars, there are two parts to the grammar – the schema that describes the document and the instance document that contains the actual data. A GML document is described using a GML Schema. This allows users and developers to describe generic geographic data sets that contain points, lines and polygons. However, the developers of GML envision communities working to define community-specific application schemas [en.wikipedia.org/wiki/GML_Application_Schemas] that are specialized extensions of GML. Using application schemas, users can refer to roads, highways, and bridges instead of points, lines and polygons. If everyone in a community agrees to use the same schemas they can exchange data easily and be sure that a road is still a road when they view it. Clients and servers with interfaces that implement the OpenGIS® Web Feature Service Interface Standard[http://www.opengeospatial.org/standards/wfs] read and write GML data. GML is also an ISO standard (ISO 19136:2007) [www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber...]. See also the GML pages on OGC Network: http://www.ogcnetwork.net/gml.				
OpenGIS GML in JPEG 2000 for Geographic Imagery Encoding Specification	1.0.0	05-047r3	Martin Kyle, David Burggraf, Sean Forde, Ron Lake	2006-01-20
The OpenGIS® GML in JPEG 2000 for Geographic Imagery Encoding Standard defines the means by which the OpenGIS® Geography Markup Language (GML) Standard [http://www.opengeospatial.org/standards/gml] is used within JPEG 2000 [www.jpeg.org/jpeg2000/] images for geographic imagery. The standard also provides packaging mechanisms for including GML within JPEG 2000 data files and specific GML application schemas to support the encoding of images within JPEG 2000 data files. JPEG 2000 is a wavelet-based image compression standard that provides the ability to include XML				

data for description of the image within the JPEG 2000 data file. See also the GML pages on OGC Network: http://www.ogcnetwork.net/gml.				
OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 1: Common architecture	1.2.1	06-103r4	John Herring	2011-05-28
The OpenGIS® Simple Features Interface Standard (SFS) provides a well-defined and common way for applications to store and access feature data in relational or object-relational databases, so that the data can be used to support other applications through a common feature model, data store and information access interface. OpenGIS Simple Features are geospatial features described using vector data elements such as points, lines and polygons.				
OpenGIS Implementation Specification for Geographic information - Simple feature access - Part 2: SQL option	1.2.1	06-104r4	John Herring	2010-08-04
The OpenGIS® Simple Features Interface Standard (SFS) provides a well-defined and common way for applications to store and access feature data in relational or object-relational databases, so that the data can be used to support other applications through a common feature model, data store and information access interface. OpenGIS Simple Features are geospatial features described using vector data elements such as points, lines and polygons.				
OpenGIS Simple Features Implementation Specification for CORBA	1.0	99-054	Peter Ladstaetter	1999-06-02
The Simple Feature Specification application programming interfaces (APIs) provide for publishing, storage, access, and simple operations on Simple Features (point, line, polygon, multi-point, etc).				
OpenGIS Styled Layer Descriptor Profile of the Web Map Service Implementation Specification	1.1.0	05-078r4	Markus Lupp	2007-08-14
The OpenGIS® Styled Layer Descriptor (SLD) Profile of the OpenGIS® Web Map Service (WMS) Encoding Standard [http://www.opengeospatial.org/standards/wms] defines an encoding that extends the WMS standard to allow user-defined symbolization and coloring of geographic feature[http://www.opengeospatial.org/ogc/glossary/f] and coverage[http://www.opengeospatial.org/ogc/glossary/c] data. SLD addresses the need for users and software to be able to control the visual portrayal of the geospatial data. The ability to define styling rules requires a styling language that the client and server can both understand. The OpenGIS® Symbology Encoding Standard (SE) [http://www.opengeospatial.org/standards/symbol] provides this language, while the SLD profile of WMS enables application of SE to WMS layers using extensions of WMS operations. Additionally, SLD defines an operation for standardized access to legend symbols.				
OpenGIS® City Geography Markup Language (CityGML) Encoding Standard	1.0	08-007r1	Gerhard Gröger, Thomas H. Kolbe, Angela Czerwinski, Claus Nagel	2008-08-20
CityGML is an open data model and XML-based format for the storage and exchange of virtual 3D city models. It is an application schema for the Geography Markup Language version 3.1.1 (GML3), the extendible international standard for spatial data exchange issued by the Open Geospatial Consortium (OGC) and the ISO TC211. The aim of the development of CityGML is to reach a common definition of the basic entities, attributes, and relations of a 3D city model. This is especially important with respect to the cost-effective sustainable maintenance of 3D city models, allowing the reuse of the same data in different application fields.				

Annex B A selection of OGC Abstract Specifications

Document Title	Version	Doc.#	Editor	Date
Topic 1 - Feature Geometry	5.0	01-101	John Herring	2001-05-10
Same as ISO 19107, available at http://www.iso.org.				
Topic 3 - Locational Geometry Structures	4.0	99-103	Cliff Kottman	1999-03-18
Provides essential and abstract models for GIS technology that is widely used.				
Topic 5 - Features	5.0	08-126	Cliff Kottman and Carl Reed	2009-01-15
From ISO 19101, “A feature is an abstraction of a real world phenomenon”; it is a geographic feature if it is associated with a location relative to the Earth.				
Topic 6 - Schema for coverage geometry and functions	7.0	07-011	OGC	2007-12-28
This International Standard defines a conceptual schema for the spatial characteristics of coverages. Coverages support mapping from a spatial, temporal or spatiotemporal domain to feature attribute values where feature attribute types are common to all geographic positions within the domain. A coverage domain consists of a collection of direct positions in a coordinate space that may be defined in terms of up to three spatial dimensions as well as a temporal dimension.				
Topic 8 - Relationships Between Features	4.0	99-108r2	Cliff Kottman	1999-03-26
This Topic introduces an abstraction for the relationships between entities in the real world. This abstraction is modeled as relationships between the features introduced in Topic 5.				
Topic 10 - Feature Collections	4.0	99-110	Cliff Kottman	1999-04-07
An OpenGIS Feature Collection is an abstract object consisting of Feature Instances, their Feature Schema, and Project Schema.				
Topic 11 - Metadata	5.0	01-111	ISO	2001-06-08
ISO 19115 was adopted as a replacement for OGC Abstract Specification Topics 9 and 11. In June 2001, a motion to include material in addition to ISO 19115 was adopted as document "01-111 Metadata AS." The approved addition to document 01-111 is contained in document 01-053r1, which normatively references parts of the old AS Topic 9, document 99-109r1. FGDC in conjunction with ANSI INCITS L1 are planning the migration of the FGDC Content Standard for Geospatial Metadata to be a profile of ISO 19115				
Topic 12 - The OpenGIS Service Architecture	4.3	02-112	ISO	2001-09-14
Same as ISO 19119				
Topic 13 - Catalog Services	4.0	99-113	Cliff Kottman	1999-03-31
Covers the Geospatial Information Access Services				
Topic 14 - Semantics and Information Communities	4.0	99-114	Cliff Kottman	1999-04-04

The OpenGIS notion of Information Communities was devised to enable groups such as ecologists and civil engineers to efficiently manage the semantics (or feature schema mismatches) of their own geodata collections and get maximum benefit from each other's geodata collections, despite semantic differences.				
Topic 15 - Image Exploitation Services	6.0	00-115	Cliff Kottman, Arliss Whiteside	2000-04-24
Describes the categories and taxonomy of image exploitation services needed to support the use of images and certain related coverage types.				

Annex C Basic schemas in GML, ISO/TC211, INSPIRE definitions

In this annex the foundational schemas in UML are reported as defined by the main standards for spatial and geographical information management: GML (and CityGML schemas are reported to, even if deriving from GML), ISO TC211 specifications, INSPIRE data model. They are related and they always pursue a reciprocal integration and adoption in order always to foster interoperability (Figure 130). They all share the same definition of the data, even if some extension can be realized in some cases, but this don't affect the conceptual nature of the structures.

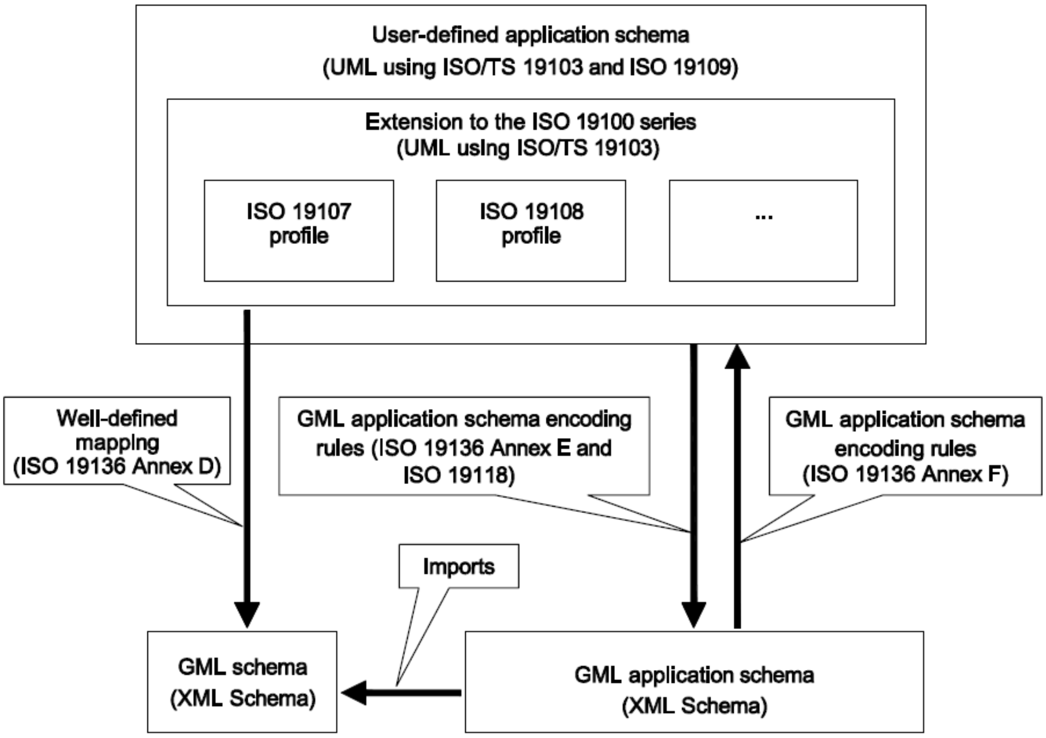
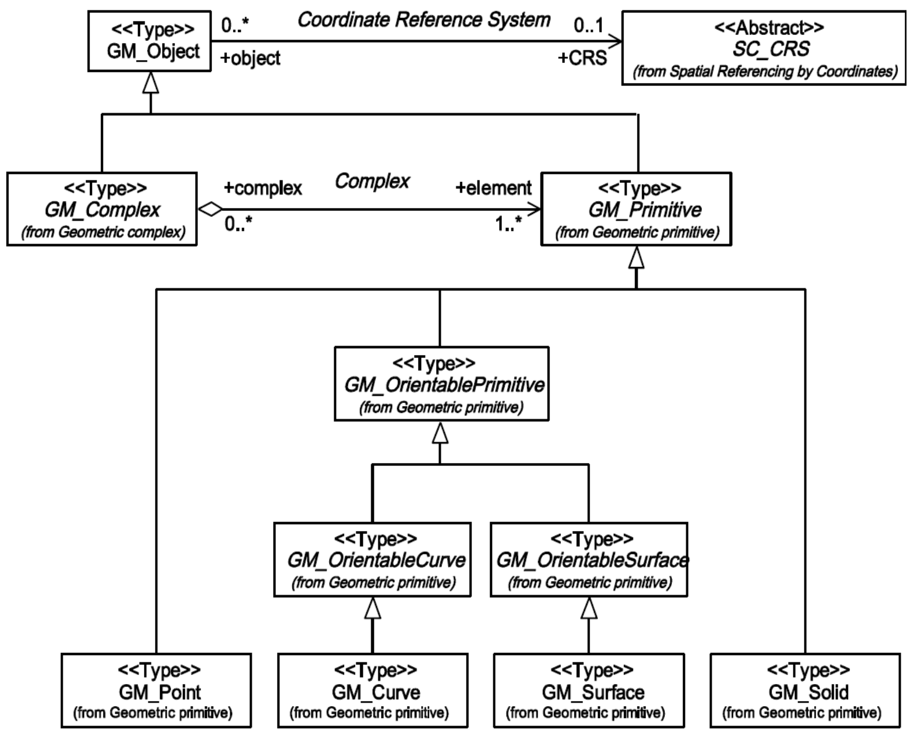


Figure 130. Schema of the relationships between the ISO 19100 series of International Standards, the GML schema and GML application schemas (OGC, 2007).

Geometric primitives



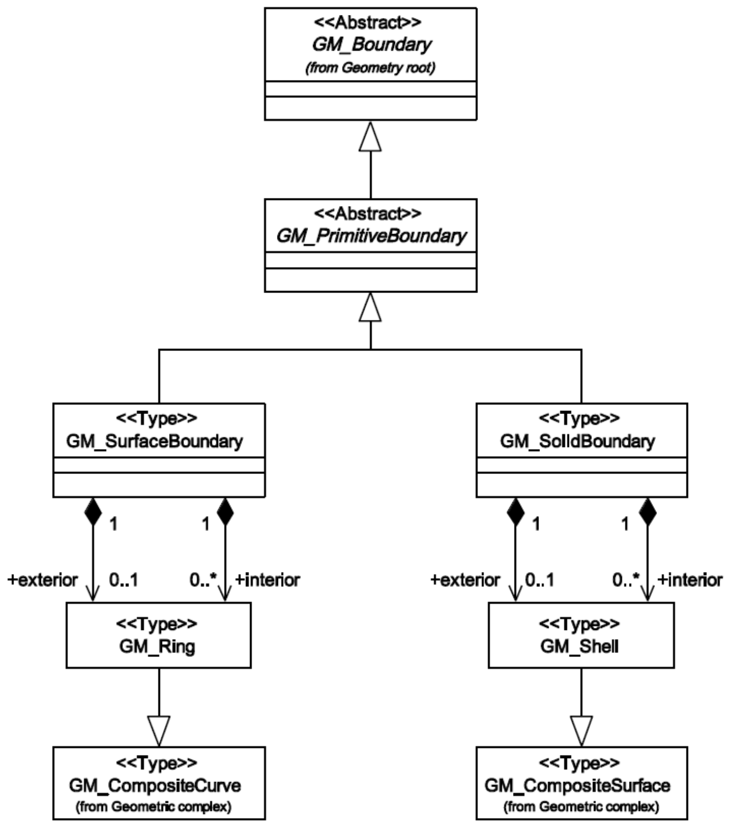


Figure 131. Geometric primitives (OGC, 2007). On the right: Boundaries of geometric primitives in GML (OGC, 2007)

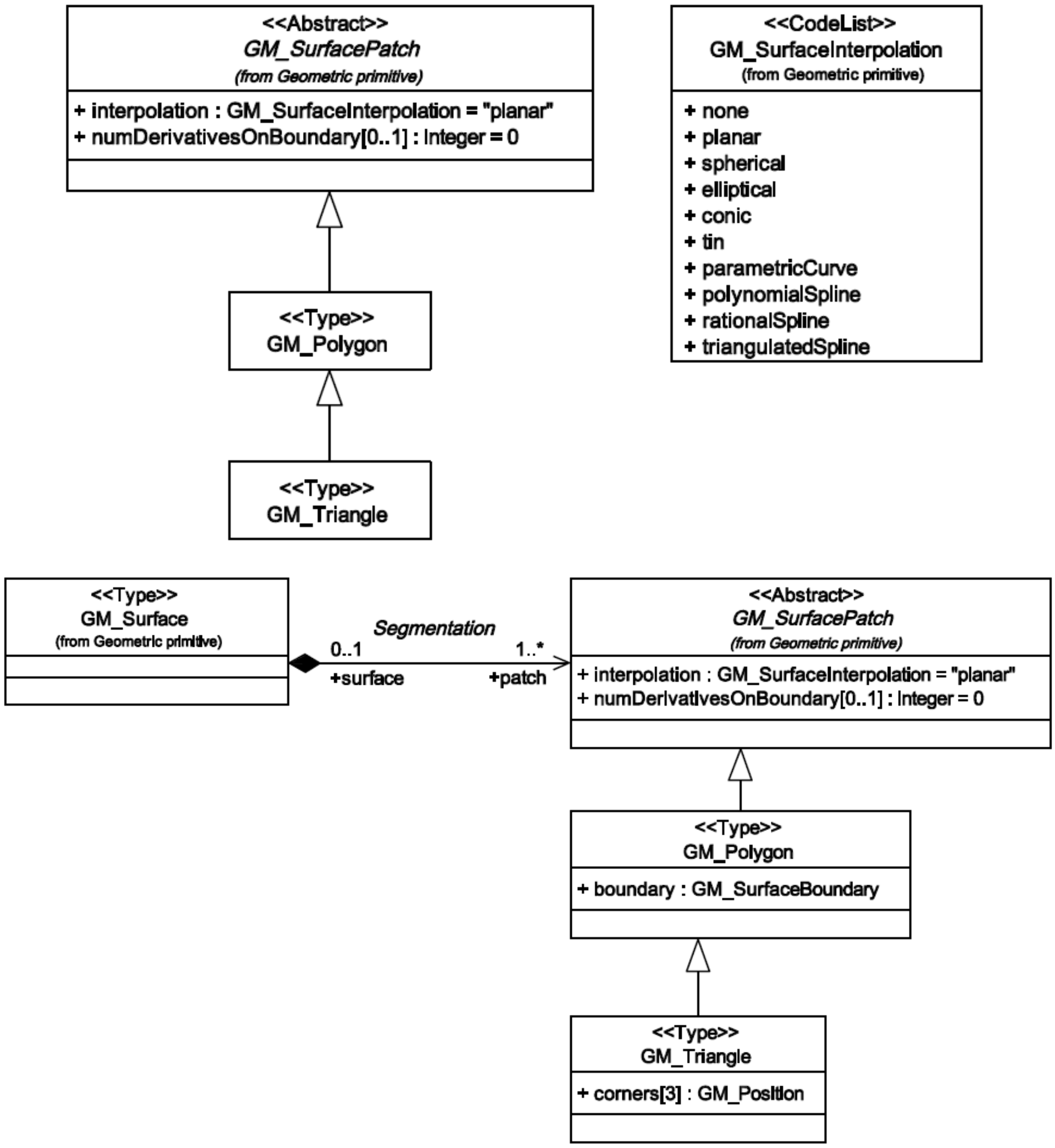


Figure 132. Surface patches I and II (OGC, 2007)

Relations between primitives for more complex objects

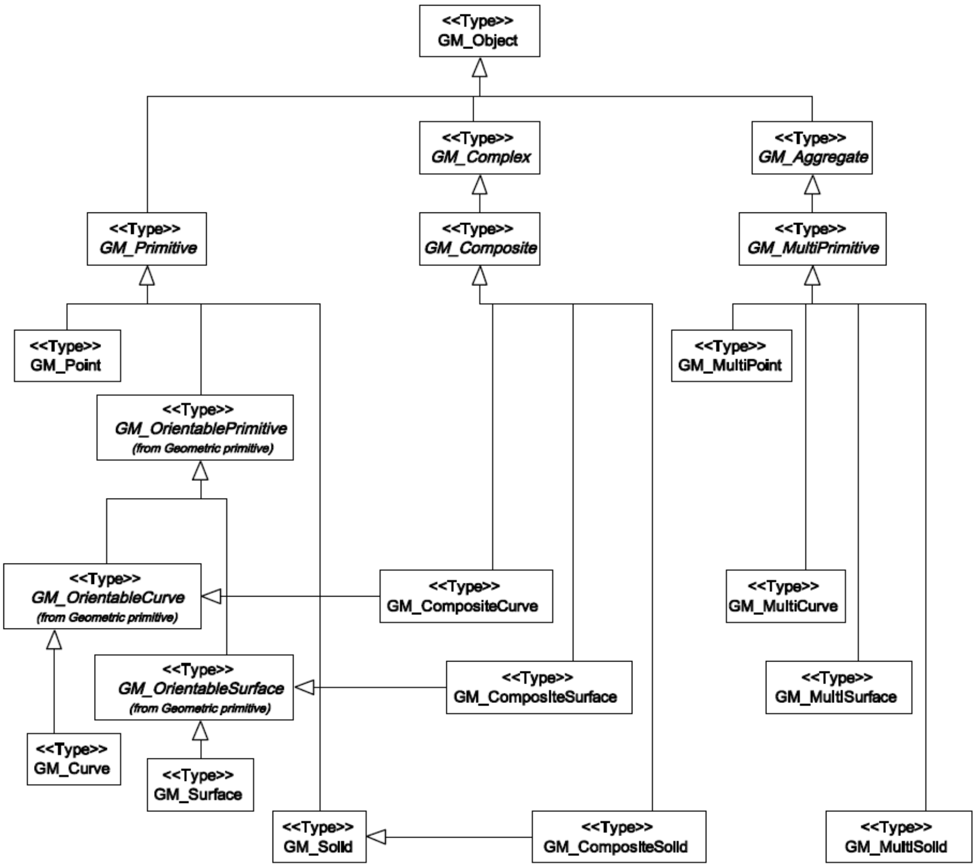


Figure 133. Implemented subtypes of GM_Object (OGC, 2007).

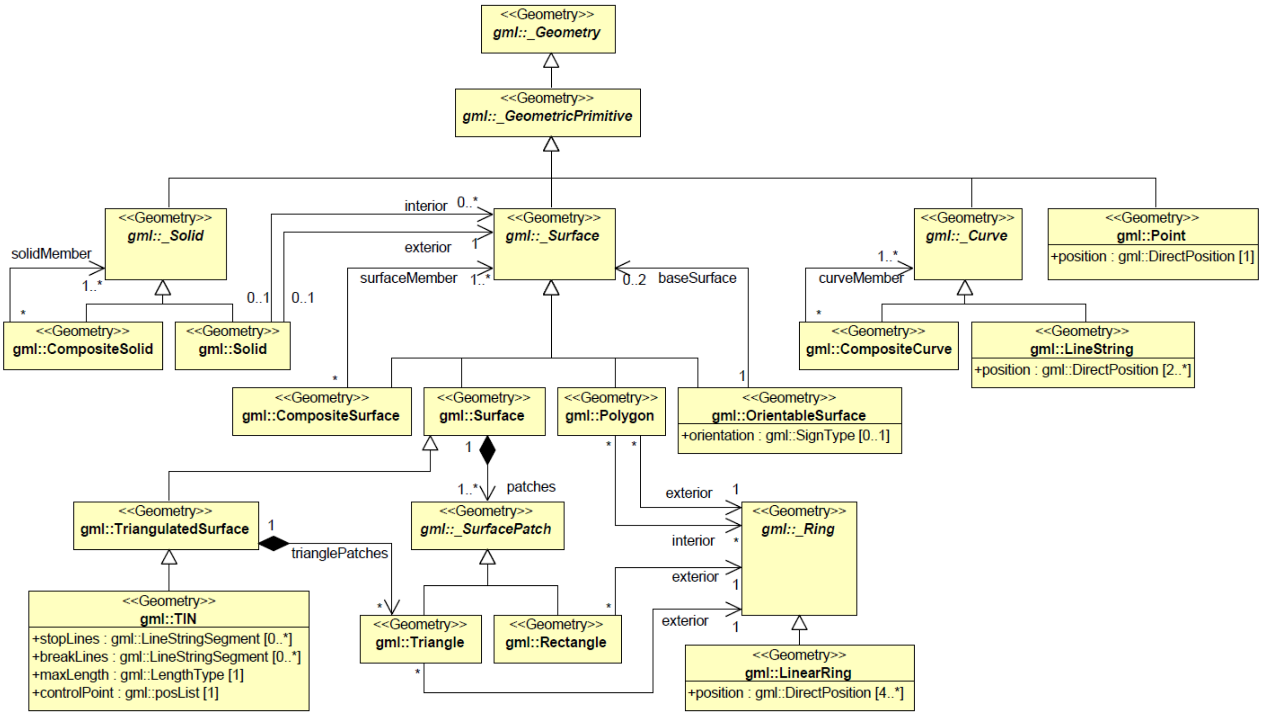


Figure 134. CityGML geometry model (subset and profile of GML3): Primitives and Composites.

Geometric composites in detail

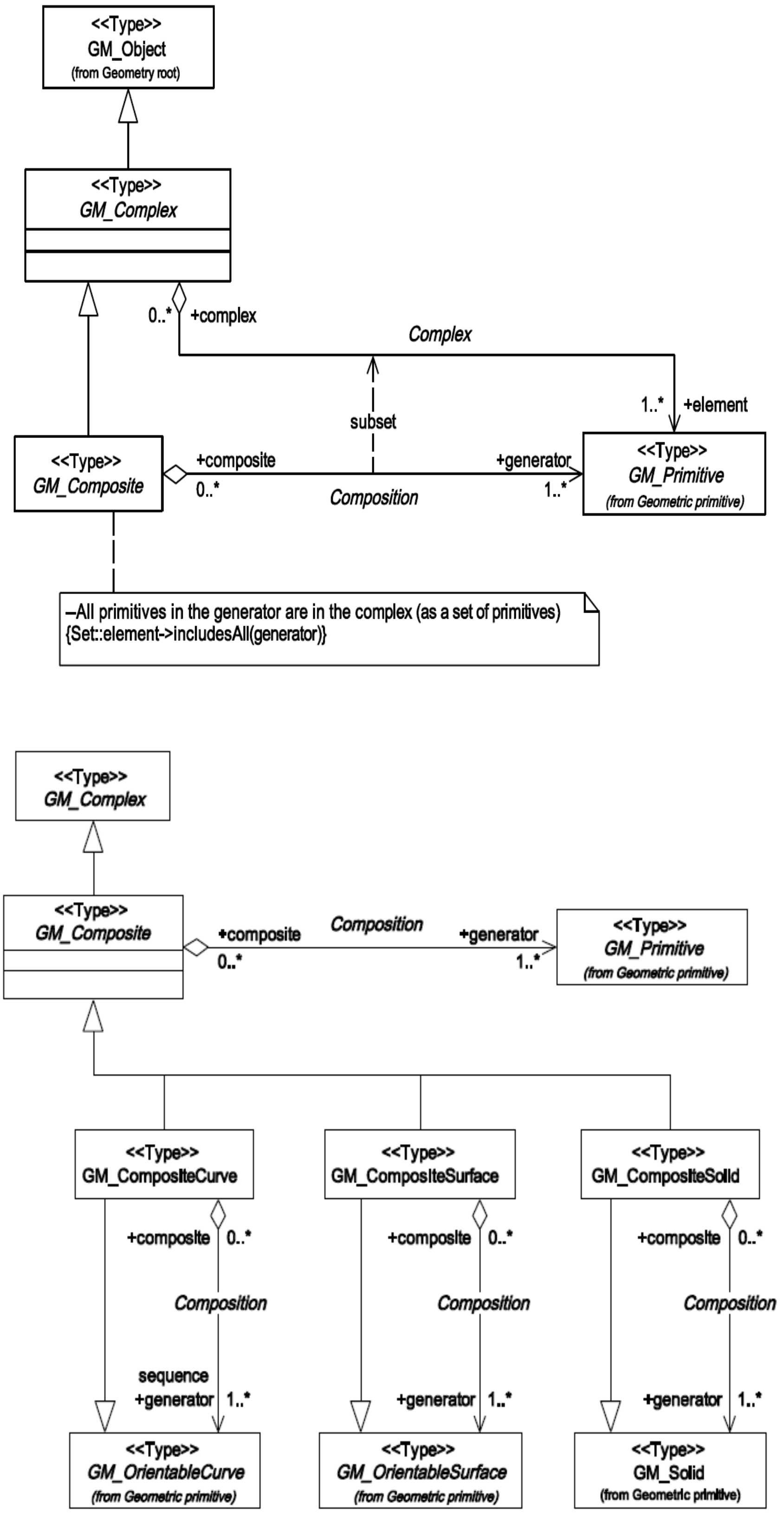


Figure 135. Geometric composites I and II (OGC, 2007). Contrary to geometric aggregates, geometric complexes can be hierarchically structured.

Topology

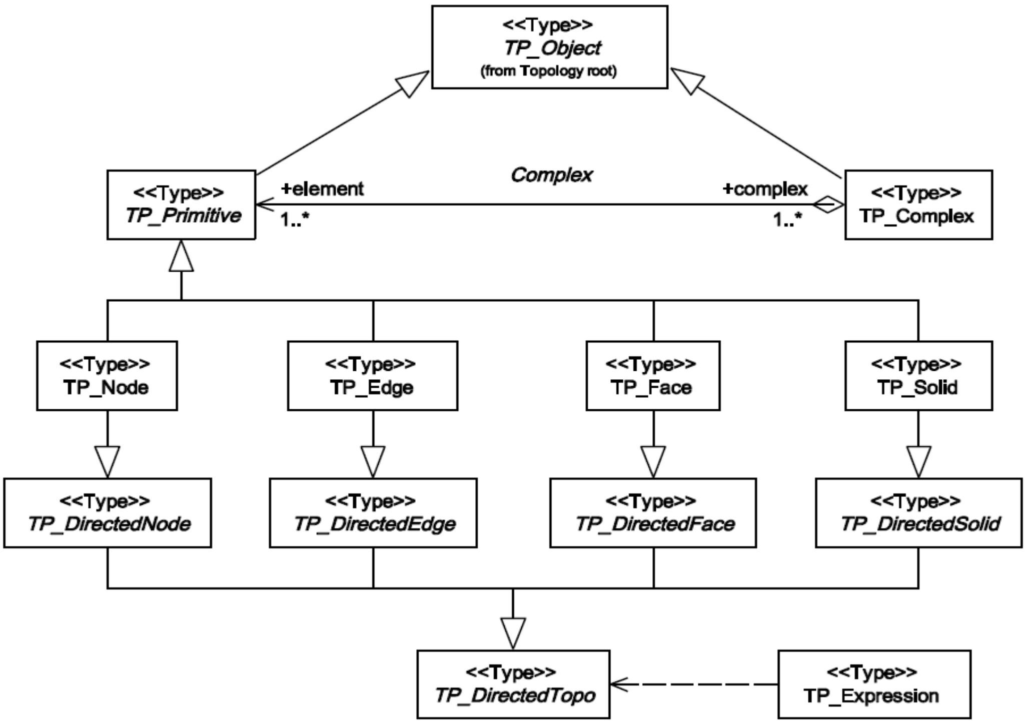


Figure 136. Topologic primitives (OGC, 2007)

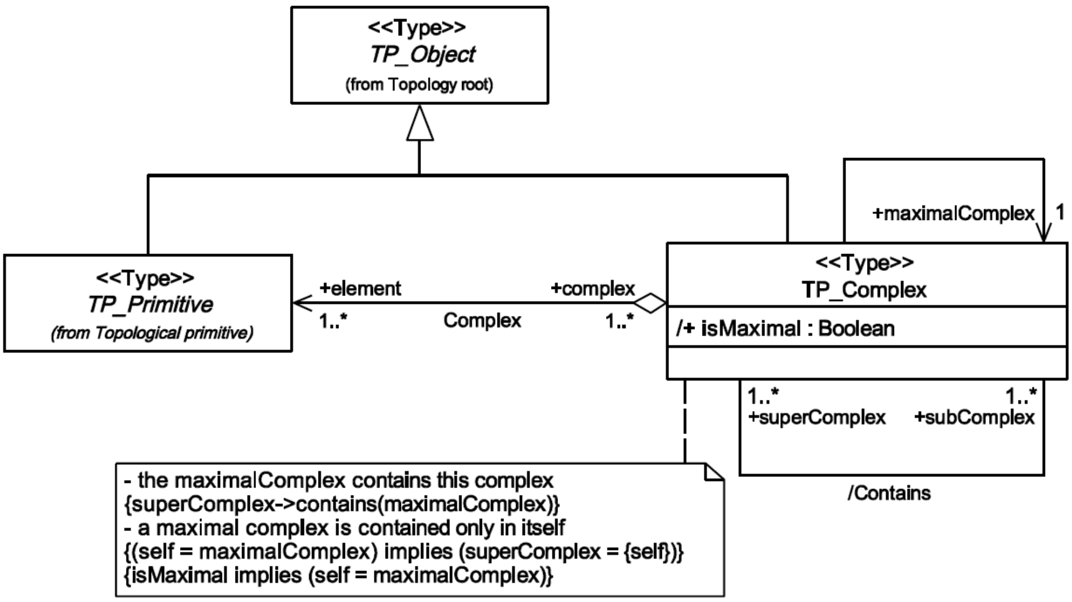


Figure 137. GML Topology complex (OGC, 2007)

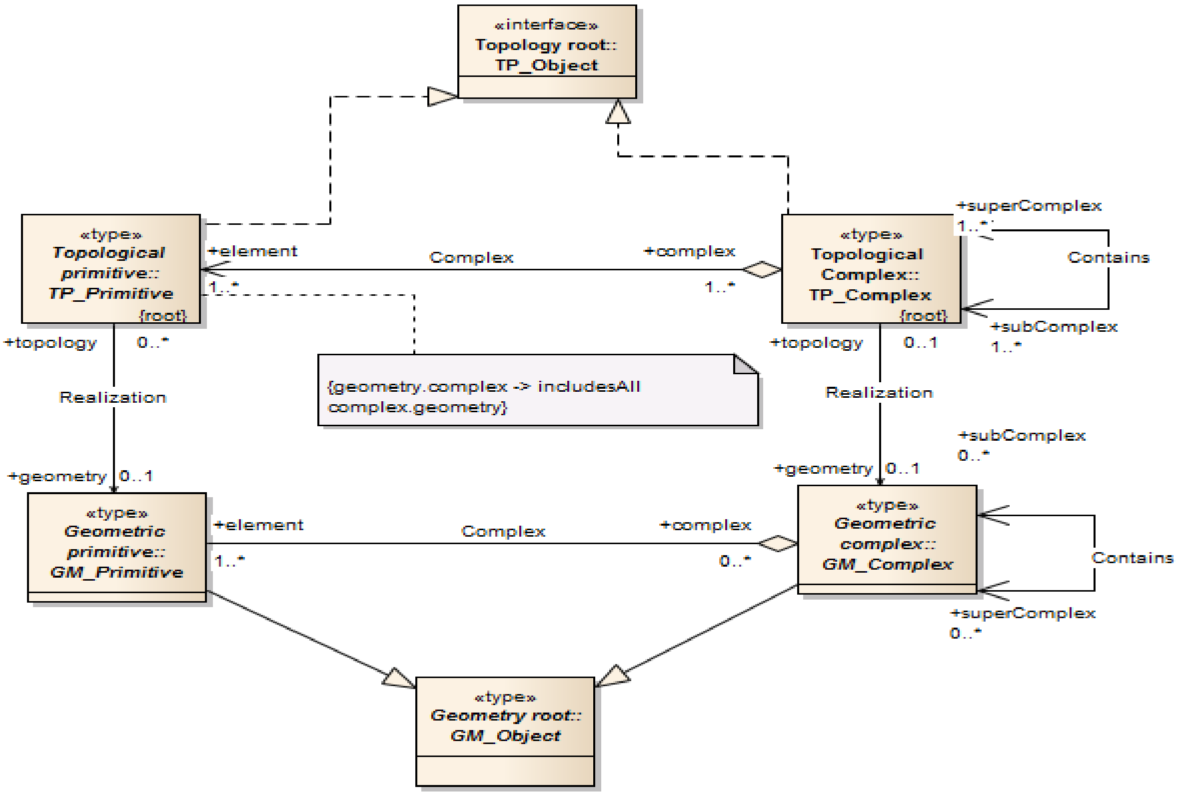
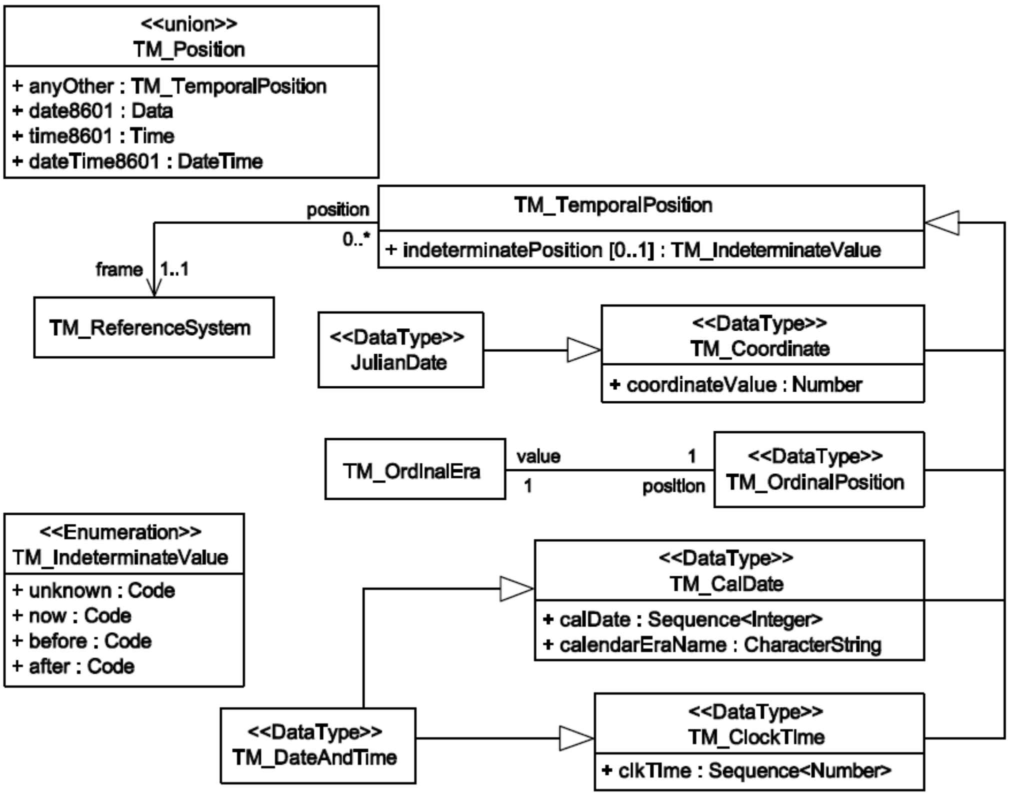
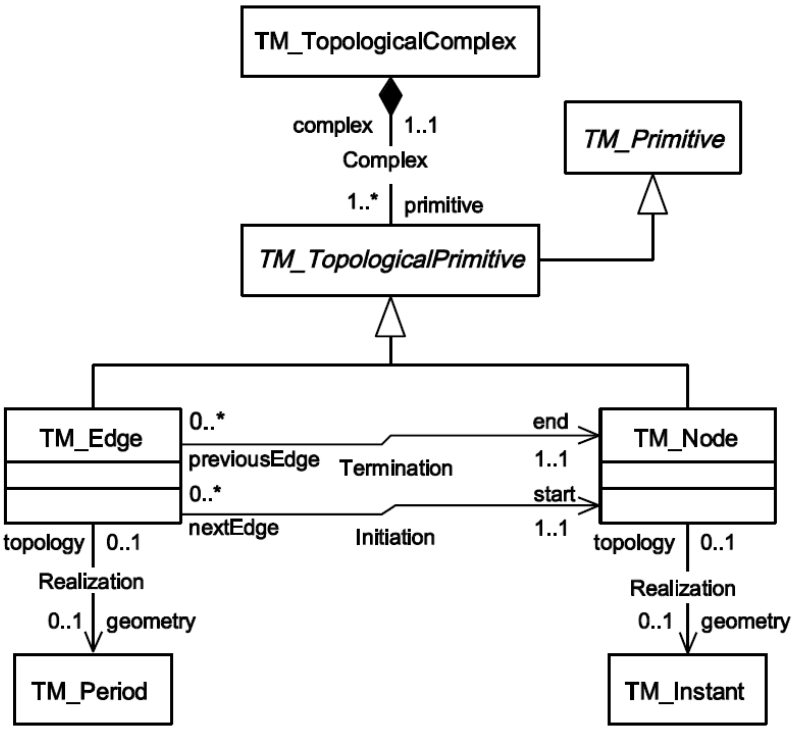


Figure 138. Relationship between geometry and topology. (http://inspire.ec.europa.eu/data-model/approved/r4618-ir/html/).

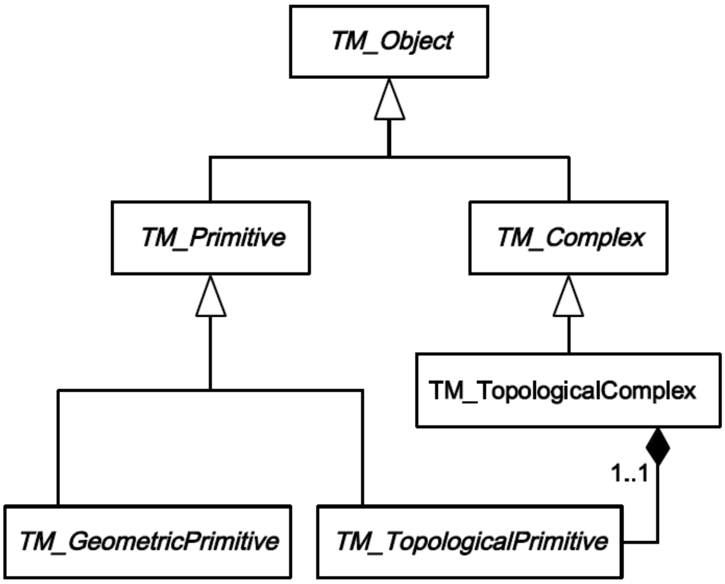


.I

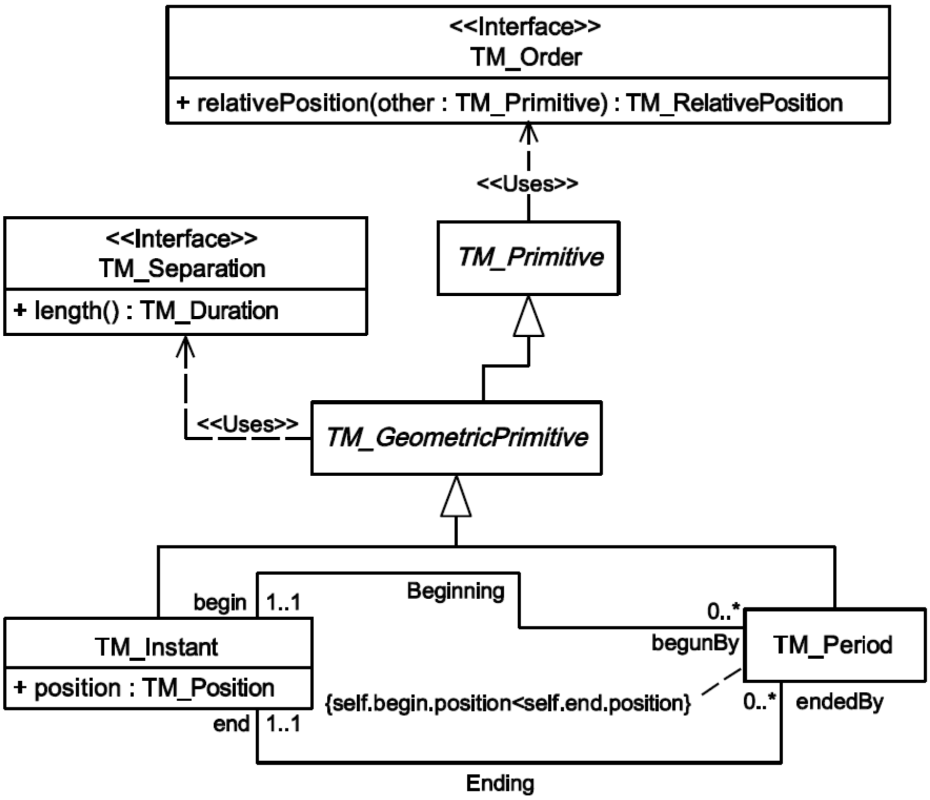


.II

Figure 139. (I) Temporal position, from ISO 19108 (OGC, 2007); (II) profile of temporal topology adapted from ISO 19108 (OGC, 2007).



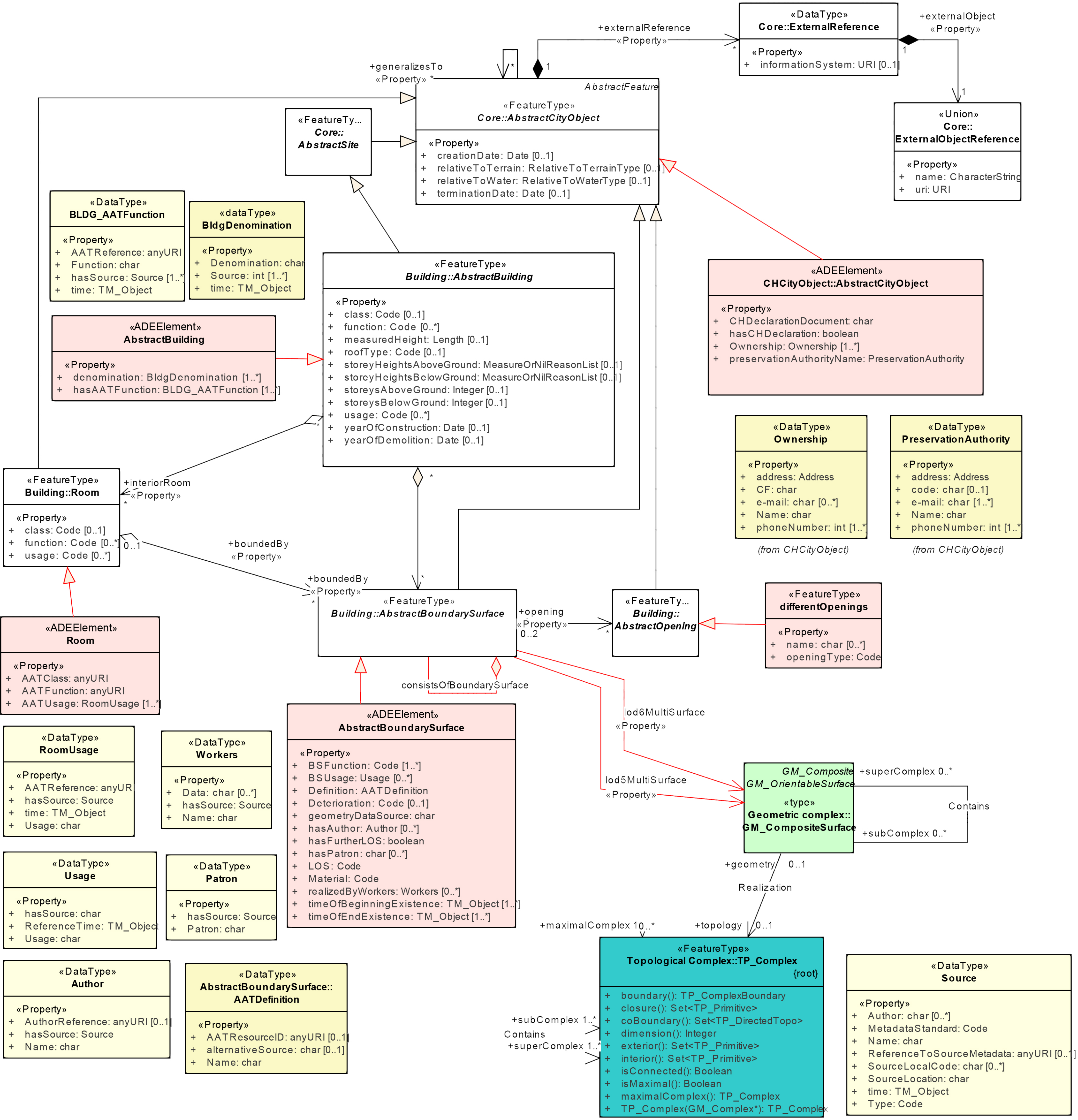
.I



.II

(I) Main hierarchy of temporal objects from ISO 19108 (OGC, 2007). (II) Pprofile of temporal geometric objects adapted from ISO 19108 (OGC, 2007)

Annex D CityGML CHADE UML model



«DataType»
Core::ExternalReference
«Property»
+ informationSystem: URI [0..1]
+externalReference
«Property»
+externalObject
«Property»
«Union»
Core::
ExternalObjectReference
«Property»
+ name: CharacterString
+ uri: URI
+generalizesTo
«Property»
AbstractFeature
«FeatureType»
Core::AbstractCityObject
«Property»
+ creationDate: Date [0..1]
+ relativeToTerrain: RelativeToTerrainType [0..1]
+ relativeToWater: RelativeToWaterType [0..1]
+ terminationDate: Date [0..1]
«FeatureTy...
Core::
AbstractSite
«DataType»
BLDG_AATFunction
«Property»
+ AATReference: anyURI
+ Function: char
+ hasSource: Source [1..*]
+ time: TM_Object
«dataType»
BldgDenomination
«Property»
+ Denomination: char
+ Source: int [1..*]
+ time: TM_Object
«FeatureType»
Building::AbstractBuilding
«Property»
+ class: Code [0..1]
+ function: Code [0..*]
+ measuredHeight: Length [0..1]
+ roofType: Code [0..1]
+ storeyHeightsAboveGround: MeasureOrNilReasonList [0..1]
+ storeyHeightsBelowGround: MeasureOrNilReasonList [0..1]
+ storeysAboveGround: Integer [0..1]
+ storeysBelowGround: Integer [0..1]
+ usage: Code [0..*]
+ yearOfConstruction: Date [0..1]
+ yearOfDemolition: Date [0..1]
«ADEElement»
CHCityObject::AbstractCityObject
«Property»
+ CHDeclarationDocument: char
+ hasCHDeclaration: boolean
+ Ownership: Ownership [1..*]
+ preservationAuthorityName: PreservationAuthority
«ADEElement»
AbstractBuilding
«Property»
+ denomination: BldgDenomination [1..*]
+ hasAATFunction: BLDG_AATFunction [1..*]
«DataType»
Ownership
«Property»
+ address: Address
+ CF: char
+ e-mail: char [0..*]
+ Name: char
+ phoneNumber: int [1..*]
(from CHCityObject)
«DataType»
PreservationAuthority
«Property»
+ address: Address
+ code: char [0..1]
+ e-mail: char [1..*]
+ Name: char
+ phoneNumber: int [1..*]
(from CHCityObject)
«FeatureType»
Building::Room
«Property»
+ class: Code [0..1]
+ function: Code [0..*]
+ usage: Code [0..*]
+interiorRoom
«Property»
+boundedBy
«Property»
0..1
+boundedBy
«Property»
«FeatureType»
Building::AbstractBoundarySurface
+opening
«Property»
0..2
«FeatureTy...
Building::
AbstractOpening
«FeatureType»
differentOpenings
«Property»
+ name: char [0..*]
+ openingType: Code
«ADEElement»
Room
«Property»
+ AATClass: anyURI
+ AATFunction: anyURI
+ AATUsage: RoomUsage [1..*]
consistsOfBoundarySurface
«ADEElement»
AbstractBoundarySurface
«Property»
+ BSFunction: Code [1..*]
+ BSUsage: Usage [0..*]
+ Definition: AATDefinition
+ Deterioration: Code [0..1]
+ geometryDataSource: char
+ hasAuthor: Author [0..*]
+ hasFurtherLOS: boolean
+ hasPatron: char [0..*]
+ LOS: Code
+ Material: Code
+ realizedByWorkers: Workers [0..*]
+ timeOfBeginningExistence: TM_Object [1..*]
+ timeOfEndExistence: TM_Object [1..*]
lod6MultiSurface
«Property»
lod5MultiSurface
«Property»
GM_Composite
GM_OrientableSurface
«type»
Geometric complex::
GM_CompositeSurface
+superComplex 0..*
Contains
+subComplex 0..*
«DataType»
RoomUsage
«Property»
+ AATReference: anyUR
+ hasSource: Source
+ time: TM_Object
+ Usage: char
«DataType»
Workers
«Property»
+ Data: char [0..*]
+ hasSource: Source
+ Name: char
«DataType»
Usage
«Property»
+ hasSource: char
+ ReferenceTime: TM_Object
+ Usage: char
«DataType»
Patron
«Property»
+ hasSource: Source
+ Patron: char
+geometry 0..1
Realization
+maximalComplex 1 0..*
+topology 0..1
«FeatureType»
Topological Complex::TP_Complex
{root}
+ boundary(): TP_ComplexBoundary
+ closure(): Set<TP_Primitive>
+ coBoundary(): Set<TP_DirectedTopo>
+ dimension(): Integer
+ exterior(): Set<TP_Primitive>
+ interior(): Set<TP_Primitive>
+ isConnected(): Boolean
+ isMaximal(): Boolean
+ maximalComplex(): TP_Complex
+ TP_Complex(GM_Complex*): TP_Complex
+subComplex 1..*
Contains
+superComplex 1..*
«DataType»
Author
«Property»
+ AuthorReference: anyURI [0..1]
+ hasSource: Source
+ Name: char
«DataType»
AbstractBoundarySurface::
AATDefinition
«Property»
+ AATResourceID: anyURI [0..1]
+ alternativeSource: char [0..1]
+ Name: char
«DataType»
Source
«Property»
+ Author: char [0..*]
+ MetadataStandard: Code
+ Name: char
+ ReferenceToSourceMetadata: anyURI [0..1]
+ SourceLocalCode: char [0..*]
+ SourceLocation: char
+ time: TM_Object
+ Type: Code

Annex E CityGML CHADE XSD schema

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!--Generated by Enterprise Architect 12.0.1215 ( Build: 1215 )-->
<xs:schema xmlns:gml="http://www.opengis.net/gml/3.2" version="1"
xmlns:ch="D:/DOTTORATO/CH_ADE" targetNamespace="D:/DOTTORATO/CH_ADE"
elementFormDefault="qualified" xmlns:xs="http://www.w3.org/2001/XMLSchema">
	<xs:include schemaLocation="CHCityObject.xsd"/>
	<xs:include schemaLocation="CHBuildings.xsd"/>
	<xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/gml.xsd"
		namespace="http://www.opengis.net/gml/3.2"/>
</xs:schema>
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSpy v2016 (x64) (http://www.altova.com) by Francesca Noardo (Politecnico di Torino) -->
<!--Generated by Enterprise Architect 12.0.1215 ( Build: 1215 )-->
<xs:schema elementFormDefault="qualified" targetNamespace="D:/DOTTORATO/CH_ADE"
xmlns:core="http://www.opengis.net/citygml/2.0" xmlns:gml="http://www.opengis.net/gml/3.2"
xmlns:ch="D:/DOTTORATO/CH_ADE" xmlns:xs="http://www.w3.org/2001/XMLSchema"
xmlns="D:/DOTTORATO/CH_ADE">
    <xs:import schemaLocation="http://schemas.opengis.net/citygml/2.0/cityGMLBase.xsd"
        namespace="http://www.opengis.net/citygml/2.0"/>
    <xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/gml.xsd"
        namespace="http://www.opengis.net/gml/3.2"/>
    <xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/gmlBase.xsd"
        namespace="http://www.opengis.net/gml/3.2"/>
    <xs:import schemaLocation="http://docs.oasis-open.org/election/external/xAL.xsd"
        namespace="urn:oasis:names:tc:ciq:xsdschema:xAL:2.0"/>
    <xs:element substitutionGroup="gml:AbstractObject" type="ch:OwnershipType"
        name="OwnershipData"/>
  - <xs:complexType name="OwnershipType">
      - <xs:sequence>
            <xs:element type="xs:int" name="phoneNumber" maxOccurs="unbounded"
                minOccurs="1"/>
            <xs:element type="xs:string" name="Name"/>
            <xs:element type="xs:string" name="e-mail" maxOccurs="unbounded"
                minOccurs="0"/>
            <xs:element type="xs:string" name="CF"/>
            <xs:element maxOccurs="unbounded" minOccurs="0" ref="core:Address"/>
        </xs:sequence>
    </xs:complexType>
  - <xs:complexType name="OwnershipPropertyType">
      - <xs:sequence>
            <xs:element ref="ch:OwnershipData"/>
        </xs:sequence>
        <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
    </xs:complexType>
    <xs:element substitutionGroup="gml:AbstractObject"
        type="ch:PreservationAuthorityType" name="PreservationAuthority"/>
  - <xs:complexType name="PreservationAuthorityType">
      - <xs:sequence>
            <xs:element type="xs:int" name="phoneNumber" maxOccurs="unbounded"
                minOccurs="1"/>
            <xs:element type="xs:string" name="Name"/>
            <xs:element type="xs:string" name="e-mail" maxOccurs="unbounded"
                minOccurs="1"/>
            <xs:element type="xs:string" name="code" maxOccurs="1" minOccurs="0"/>
            <xs:element maxOccurs="unbounded" minOccurs="0" ref="core:Address"/>
        </xs:sequence>
    </xs:complexType>
  - <xs:complexType name="PreservationAuthorityPropertyType">
      - <xs:sequence>
            <xs:element ref="ch:PreservationAuthority"/>
        </xs:sequence>
        <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
```

```
    </xs:complexType>
      <!--=========== Application specific attributes for AbstractCityObject
      ==============-->
  - <xs:element substitutionGroup="core:_GenericApplicationPropertyOfCityObject"
    name="preservationAuthorityName">
    - <xs:complexType>
      - <xs:complexContent>
        - <xs:extension base="gml:AbstractMemberType">
          - <xs:sequence>
              <xs:element ref="ch:PreservationAuthority"/>
            </xs:sequence>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
  - <xs:element substitutionGroup="core:_GenericApplicationPropertyOfCityObject"
    name="Ownership">
    - <xs:complexType>
      - <xs:complexContent>
        - <xs:extension base="gml:AbstractMemberType">
          - <xs:sequence>
              <xs:element maxOccurs="unbounded" ref="ch:OwnershipData"/>
            </xs:sequence>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
    <xs:element substitutionGroup="core:_GenericApplicationPropertyOfCityObject"
      type="xs:boolean" name="hasCHDeclaration"/>
    <xs:element substitutionGroup="core:_GenericApplicationPropertyOfCityObject"
      type="xs:string" name="CHDeclarationDocument"/>
</xs:schema>
```

```
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XMLSpy v2016 (x64) (http://www.altova.com) by Francesca Noardo (Politecnico di
Torino) -->
<!--Generated by Enterprise Architect 12.0.1215 ( Build: 1215 )-->
<xs:schema elementFormDefault="qualified" targetNamespace="D:/DOTTORATO/CH_ADE"
xmlns:bldg="http://www.opengis.net/citygml/building/2.0"
xmlns:gml="http://www.opengis.net/gml/3.2" xmlns:ch="D:/DOTTORATO/CH_ADE"
xmlns:core="http://www.opengis.net/citygml/2.0"
xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns="D:/DOTTORATO/CH_ADE">
    <xs:import schemaLocation="http://schemas.opengis.net/citygml/building/2.0/building.xsd"
        namespace="http://www.opengis.net/citygml/building/2.0"/>
    <xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/gml.xsd"
        namespace="http://www.opengis.net/gml/3.2"/>
    <xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/gmlBase.xsd"
        namespace="http://www.opengis.net/gml/3.2"/>
    <xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/temporal.xsd"
        namespace="http://www.opengis.net/gml/3.2"/>
    <xs:import schemaLocation="http://schemas.opengis.net/gml/3.2.1/basicTypes.xsd"
        namespace="http://www.opengis.net/gml/3.2"/>
        <!--======== Application specific attributes for AbstractBoundarySurface
        =============-->
  - <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
    type="WorkersType" name="realizedByWorkers">
    - <xs:annotation>
            <xs:documentation>Reference to existing vocabularies, when
                possible</xs:documentation>
        </xs:annotation>
    </xs:element>
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
        type="gml:AbstractTimeObjectType" name="timeOfBeginningExistence"/>
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
        type="gml:AbstractTimeObjectType" name="timeOfEndExistence"/>
  - <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
    type="xs:anyURI" name="Material">
    - <xs:annotation>
            <xs:documentation>Reference to CIDOC-MONDIS owl
                ontology</xs:documentation>
        </xs:annotation>
    </xs:element>
  - <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
    type="xs:anyURI" name="Deterioration">
    - <xs:annotation>
            <xs:documentation>Reference to CIDOC-MONDIS owl
                ontology</xs:documentation>
        </xs:annotation>
    </xs:element>
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
        type="gml:CodeType" name="LOS"/>
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
        type="xs:boolean" name="hasFurtherLOS"/>
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
        type="AuthorType" name="hasAuthor"/>
```

```
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
      type="PatronType" name="hasPatron"/>
- <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
  type="xs:string" name="geometryDataSource">
  - <xs:annotation>
      <xs:documentation>It should refer to a chosen format of metadata for the survey and
        the cartographic product</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
      type="AATDefinitionType" name="Definition"/>
- <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
  type="UsageType" name="BSUsage">
  - <xs:annotation>
      <xs:documentation>The usage of the element, independently from its original
        function. To be specified if they are different. It's recommended to refer to the
        AAT Vocabulary of the Getty Institute
        (http://vocab.getty.edu/).</xs:documentation>
    </xs:annotation>
  </xs:element>
- <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
  type="xs:string" name="BSFunction">
  - <xs:annotation>
      <xs:documentation>Function of the element or the part of element to which the
        surface belongs to. The codelist refers to the AAT Vocabulary of the Getty
        Institute (http://vocab.getty.edu/).</xs:documentation>
    </xs:annotation>
  </xs:element>
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
      type="gml:MultiSurfacePropertyType" name="lod5MultiSurface"/>
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBoundarySurface"
      type="gml:MultiSurfacePropertyType" name="lod6MultiSurface"/>
      <!--======== Application specific attributes for AbstractBuilding ============-->
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBuilding"
      type="BLDG_AATFunctionType" name="hasAATFunction"/>
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBuilding"
      type="BldgDenominationType" name="denomination"/>
      <!--======== New classes and data types ============-->
  <xs:element substitutionGroup="gml:AbstractObject" type="AuthorType"
      name="Author"/>
- <xs:complexType name="AuthorType">
  - <xs:sequence>
      <xs:element type="xs:string" name="Name"/>
      <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
        minOccurs="1"/>
    - <xs:element type="xs:anyURI" name="AuthorReference" maxOccurs="1"
      minOccurs="0">
      - <xs:annotation>
          <xs:documentation>Reference to Getty Vocabulary ULAN (Union List od
            Artist Names) in the form "http://vocab.getty.edu/ulan/500022389" It
            is the resource name in the Getty Vocabulary</xs:documentation>
        </xs:annotation>
      </xs:element>
```

```
        </xs:sequence>
    </xs:complexType>
 - <xs:complexType name="AuthorPropertyType">
    - <xs:sequence>
          <xs:element ref="Author"/>
      </xs:sequence>
      <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
   </xs:complexType>
   <xs:element substitutionGroup="gml:AbstractObject" type="AATDefinitionType"
      name="AATDefinition"/>
 - <xs:complexType name="AATDefinitionType">
    - <xs:sequence>
       - <xs:element type="xs:anyURI" name="AATResource" maxOccurs="1"
          minOccurs="0">
          - <xs:annotation>
                <xs:documentation>Reference to Getty Vocabulary
                   AAT</xs:documentation>
             </xs:annotation>
          </xs:element>
          <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
             minOccurs="1"/>
          <xs:element type="xs:string" name="Name"/>
      </xs:sequence>
   </xs:complexType>
 - <xs:complexType name="AATDefinitionPropertyType">
    - <xs:sequence>
          <xs:element ref="AATDefinition"/>
      </xs:sequence>
      <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
   </xs:complexType>
   <xs:element substitutionGroup="gml:AbstractObject" type="BLDG_AATFunctionType"
      name="BLDG_AATFunction"/>
 - <xs:complexType name="BLDG_AATFunctionType">
    - <xs:sequence>
          <xs:element type="gml:AbstractTimeObjectType" name="time"/>
          <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
             minOccurs="1"/>
          <xs:element type="xs:string" name="Function"/>
       - <xs:element type="xs:anyURI" name="AATReference">
          - <xs:annotation>
                <xs:documentation>It is defined by the AAT vocabulary of the Getty
                   Institute, as a subclass of <single built works by function>
                   (http://vocab.getty.edu/aat/300004894) in the form:
                   http://vocab.getty.edu/aat/'numberID'</xs:documentation>
             </xs:annotation>
          </xs:element>
      </xs:sequence>
   </xs:complexType>
 - <xs:complexType name="BLDG_AATFunctionPropertyType">
    - <xs:sequence>
          <xs:element ref="BLDG_AATFunction"/>
      </xs:sequence>
      <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
```

```
    </xs:complexType>
    <xs:element substitutionGroup="gml:AbstractObject" type="PatronType" name="Patron"/>
  - <xs:complexType name="PatronType">
      - <xs:sequence>
          - <xs:element type="xs:string" name="Patron">
              - <xs:annotation>
                    <xs:documentation>If possible the reference to some vocabulary will be
                        encouraged in future developments.</xs:documentation>
                </xs:annotation>
            </xs:element>
            <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
                minOccurs="1"/>
        </xs:sequence>
    </xs:complexType>
  - <xs:complexType name="PatronPropertyType">
      - <xs:sequence>
            <xs:element ref="Patron"/>
        </xs:sequence>
        <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
    </xs:complexType>
        <!--======== Application specific attributes for Room ============-->
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfRoom"
        type="RoomUsageType" name="AATUsage"/>
  - <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfAbstractBuilding"
    type="xs:anyURI" name="AATFunction">
      - <xs:annotation>
            <xs:documentation>It is defined by the AAT vocabulary of the Getty Institute, as a
                subclass of <room and spaces by function>
                (http://vocab.getty.edu/aat/300004021) in the form:
                http://vocab.getty.edu/aat/'numberID'</xs:documentation>
        </xs:annotation>
    </xs:element>
  - <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfAbstractBuilding"
    type="xs:anyURI" name="AATClass">
      - <xs:annotation>
            <xs:documentation>It is defined by the AAT vocabulary of the Getty Institute, as a
                subclass of 'room and spaces by form'; (http://vocab.getty.edu/aat/300004053) or
                subclass of 'room and spaces by location or context';
                (http://vocab.getty.edu/aat/300004011) subclass of 'room and spaces by building
                type'; (http://vocab.getty.edu/aat/30010051) in the form:
                http://vocab.getty.edu/aat/'numberID'</xs:documentation>
        </xs:annotation>
    </xs:element>
        <!--======== Application specific attributes for AbstractOpenings ============-->
    <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfOpening"
        type="xs:anyURI" name="AATOpeningReference"/>
        <!--======== New classes and data types ============-->
    <xs:element substitutionGroup="gml:AbstractObject" type="RoomUsageType"
        name="RoomUsage"/>
  - <xs:complexType name="RoomUsageType">
      - <xs:sequence>
            <xs:element type="xs:string" name="Usage"/>
            <xs:element type="gml:AbstractTimeObjectType" name="time"/>
```

```
      <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
        minOccurs="1"/>
    - <xs:element type="xs:anyURI" name="AATReference">
      - <xs:annotation>
          <xs:documentation>It is defined by the AAT vocabulary of the Getty
            Institute, as a subclass of <room and spaces by function>
            (http://vocab.getty.edu/aat/300004021) in the form:
            http://vocab.getty.edu/aat/'numberID'</xs:documentation>
        </xs:annotation>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
- <xs:complexType name="RoomUsagePropertyType">
  - <xs:sequence>
      <xs:element ref="RoomUsage"/>
    </xs:sequence>
    <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
  </xs:complexType>
  <xs:element substitutionGroup="gml:AbstractObject" type="SourceType" name="Source"/>
- <xs:complexType name="SourceType">
  - <xs:sequence>
      <xs:element type="gml:CodeType" name="Type"/>
      <xs:element type="gml:AbstractTimeObjectType" name="time"/>
      <xs:element type="xs:string" name="SourceLocation"/>
    - <xs:element type="xs:string" name="SourceLocalCode" maxOccurs="unbounded"
      minOccurs="0">
      - <xs:annotation>
          <xs:documentation>e.g. Code of the document in the
            archive</xs:documentation>
        </xs:annotation>
      </xs:element>
      <xs:element type="xs:anyURI" name="ReferenceToSourceMetadata"
        maxOccurs="1" minOccurs="0"/>
      <xs:element type="xs:string" name="Name"/>
      <xs:element type="gml:CodeType" name="MetadataStandard"/>
      <xs:element type="xs:string" name="Author" maxOccurs="unbounded"
        minOccurs="0"/>
    </xs:sequence>
  </xs:complexType>
- <xs:complexType name="SourcePropertyType">
  - <xs:sequence>
      <xs:element ref="Source"/>
    </xs:sequence>
    <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
  </xs:complexType>
  <xs:element substitutionGroup="gml:AbstractObject" type="UsageType" name="Usage"/>
- <xs:complexType name="UsageType">
  - <xs:sequence>
    - <xs:element type="xs:string" name="Usage">
      - <xs:annotation>
          <xs:documentation>may be a codelist value or reference to existing
            vocabularies</xs:documentation>
        </xs:annotation>
```

```
          </xs:element>
          <xs:element type="gml:AbstractTimeObjectType" name="ReferenceTime"/>
          <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
            minOccurs="1"/>
      </xs:sequence>
  </xs:complexType>
- <xs:complexType name="UsagePropertyType">
  - <xs:sequence>
      <xs:element ref="Usage"/>
    </xs:sequence>
    <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
  </xs:complexType>
  <xs:element substitutionGroup="gml:AbstractObject" type="WorkersType"
    name="Workers"/>
- <xs:complexType name="WorkersType">
  - <xs:sequence>
      <xs:element type="xs:string" name="Name"/>
      <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
        minOccurs="1"/>
    - <xs:element type="xs:string" name="Data" maxOccurs="unbounded"
      minOccurs="0">
      - <xs:annotation>
          <xs:documentation>Data about the workers (contact data, address data, in
            case of known workers, or anything which could be of interest for
            defining them in the past)</xs:documentation>
        </xs:annotation>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
- <xs:complexType name="WorkersPropertyType">
  - <xs:sequence>
      <xs:element ref="Workers"/>
    </xs:sequence>
    <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
  </xs:complexType>
  <xs:element substitutionGroup="bldg:_GenericApplicationPropertyOfBuilding"
    type="BldgDenominationType" name="BldgDenomination"/>
- <xs:complexType name="BldgDenominationPropertyType">
  - <xs:sequence minOccurs="0">
      <xs:element ref="BldgDenomination"/>
    </xs:sequence>
    <xs:attributeGroup ref="gml:AssociationAttributeGroup"/>
    <xs:attributeGroup ref="gml:OwnershipAttributeGroup"/>
  </xs:complexType>
- <xs:complexType name="BldgDenominationType">
  - <xs:sequence>
      <xs:element type="gml:AbstractTimeObjectType" name="time"/>
      <xs:element type="SourceType" name="hasSource" maxOccurs="unbounded"
        minOccurs="1"/>
      <xs:element type="xs:string" name="Denomination"/>
    </xs:sequence>
  </xs:complexType>
</xs:schema>
```